



**UNIVERSIDAD TECNOLÓGICA CENTROAMERICANA**

**FACULTAD DE INGENIERÍA**

**PROYECTO DE INVESTIGACIÓN**

**ESTUDIO COMPARATIVO DE MODELOS DE INTELIGENCIA ARTIFICIAL PARA LA  
IDENTIFICACIÓN DE TUMORES EN ECOGRAFÍAS MAMOGRÁFICAS**

**PREVIO A LA OBTENCIÓN DEL TÍTULO:**

**INGENIERO EN BIOMÉDICA**

**PRESENTADO POR:**

**22011150 KENNETH JOSUE LAGOS MOTIÑO**

**ASESOR: MANUEL GAMERO**

**CAMPUS: SAN PEDRO SULA**

**JUNIO, 2024**

## **DEDICATORIA**

Dedico este trabajo de investigación principalmente a Dios, quien me guió en todo el camino. A mis padres, Julio Lagos y Leila Motiño, quienes son las personas que más admiro. Agradezco profundamente cada esfuerzo que han hecho por mí, sus consejos, su amor, sus enseñanzas y por haberme formado a lo largo de toda esta trayectoria. Gracias por confiar en mí y por darme esta oportunidad. A mi hermano, amigos y mi familia que me ha apoyado.

## **AGRADECIMIENTOS**

Agradezco principalmente a Dios por la oportunidad de culminar este trabajo.

A mi padre Julio Lagos por todo el esfuerzo que ha hecho, por darme todo su apoyo, su amor incondicional para culminar esta etapa.

Al Ing Manuel Gamero, quien me ha guiado durante este trabajo de investigación, por sus consejos, su paciencia y por dedicarme tiempo durante este periodo académico.

A la Abogada Karla Aguilar por darme su asesoría y su tiempo para poder culminar este trabajo.

## RESUMEN EJECUTIVO

El presente estudio ha evaluado diferentes modelos de inteligencia artificial para la clasificación de tumores en imágenes de ecografías mamarias. El objetivo general ha sido evaluar y comparar el rendimiento de diversas arquitecturas de redes neuronales convolucionales, implementando técnicas de preprocesamiento y validación para mejorar la precisión en la identificación de tumores.

Se ha seleccionado una base de datos diversificada de imágenes de ecografías mamarias, asegurando una clasificación adecuada en normal, benigno y maligno. Este proceso ha incluido técnicas de preprocesamiento y aumentación de datos para mejorar la robustez del modelo de inteligencia artificial. Se ha analizado la influencia del tamaño del lote en la precisión y pérdida del modelo durante su entrenamiento, identificando que un tamaño de lote más pequeño ha proporcionado una mayor precisión y estabilidad.

Además, se han comparado el rendimiento de diferentes arquitecturas de redes neuronales convolucionales, como EfficientNetV2, ResNet50 y VGG16, utilizando un conjunto de datos de prueba y métricas como la precisión, la pérdida y el tiempo de entrenamiento. EfficientNetV2 ha demostrado ser el modelo más eficiente y preciso para la clasificación de tumores mamarios. Finalmente, se ha validado el modelo de inteligencia artificial utilizando un conjunto de datos independiente, asegurando su fiabilidad y capacidad de generalización en condiciones clínicas reales.

Los principales resultados han indicado que la elección del modelo EfficientNetV2 ha sido adecuada, alcanzando una precisión del 95% en la identificación de tumores. Las conclusiones han destacado el potencial del uso de inteligencia artificial en la detección temprana del cáncer de mama, sugiriendo su implementación como una herramienta de apoyo en el diagnóstico médico.

Palabras clave: Clasificación, ecografía mamaria, inteligencia artificial, preprocesamiento imágenes, redes neuronales, tumores.

## **ABSTRACT**

The present study has evaluated different artificial intelligence models for the detection of tumors in breast ultrasound images. The general objective has been to evaluate and compare the performance of various convolutional neural network architectures, implementing preprocessing and validation techniques to improve the accuracy in tumor identification.

A diversified database of breast ultrasound images has been selected, ensuring an adequate classification into normal, benign and malignant. This process has included preprocessing and data augmentation techniques to improve the robustness of the artificial intelligence model. The influence of the batch size on the precision and loss of the model during its training has been analyzed, identifying that a smaller batch size has provided greater precision and stability.

Furthermore, the performance of different convolutional neural network architectures, such as EfficientNetV2, ResNet50 and VGG16, have been compared using a test data set and metrics such as accuracy, loss and training time. EfficientNetV2 has proven to be the most efficient and accurate model for breast tumor classification. Finally, the artificial intelligence model has been validated using an independent data set, ensuring its reliability and generalization capacity in real clinical conditions.

The main results have indicated that the choice of the EfficientNetV2 model has been appropriate, reaching an accuracy of 95% in tumor identification. The conclusions have highlighted the potential of the use of artificial intelligence in the early detection of breast cancer, suggesting its implementation as a support tool in medical diagnosis.

Keywords: artificial intelligence, convolutional neural, breast ultrasound, tumor detection, image preprocessing

# ÍNDICE CONTENIDO

|  |    |
|--|----|
| <b>I. INTRODUCCIÓN</b> .....                   | 3  |
| <b>II. ESTADO DEL ARTE</b> .....               | 5  |
| 2.1. ANTECEDENTES .....                        | 5  |
| 2.2. PROBLEMÁTICA .....                        | 16 |
| 2.3. IMAGEN INTEGRADORA.....                   | 18 |
| 2.4. TABLA DE LIMITACIONES.....                | 19 |
| <b>III. OBJETIVOS</b> .....                    | 24 |
| 3.1. OBJETIVO GENERAL.....                     | 24 |
| 3.2. OBJETIVOS ESPECÍFICOS:.....               | 24 |
| <b>IV. MÉTODOS</b> .....                       | 25 |
| 4.1. ENFOQUE .....                             | 25 |
| 4.2. OBJETIVO DEL MÉTODO .....                 | 25 |
| 4.3. VARIABLES DE INVESTIGACIÓN.....           | 27 |
| 4.3.1. VARIABLE DEPENDIENTES.....              | 27 |
| 4.3.2. VARIABLES INDEPENDIENTES.....           | 27 |
| 4.4. RESUMEN DEL PROCEDIMIENTO .....           | 29 |
| 4.4.1. POBLACIÓN Y MUESTRA.....                | 29 |
| 4.4.2. TÉCNICAS E INSTRUMENTOS APLICADOS ..... | 29 |
| 4.4.3. METODOLOGÍA DEL ESTUDIO.....            | 30 |
| 4.5. IMAGEN DEL PROCESO .....                  | 32 |
| 4.6. MÉTRICAS DE EVALUACIÓN Y VALIDACIÓN.....  | 33 |
| 4.6.1. PRECISIÓN (ACCURACY) .....              | 33 |
| 4.6.2. RECALL, O SENSIBILIDAD .....            | 34 |
| 4.6.3. F1 SCORE.....                           | 34 |
| 4.6.4. MATRIZ DE CONFUSIÓN .....               | 34 |
| 4.7. CRONOGRAMA DE ACTIVIDADES.....            | 35 |
| <b>V. RESULTADOS</b> .....                     | 36 |
| 5.1. BASE DE DATOS.....                        | 36 |
| 5.2. MODELO DE APRENDIZAJE.....                | 36 |

|               |  |           |
|---------------|--|-----------|
| 5.3.          | PROCESAMIENTO Y PREPARACIÓN DE IMÁGENES .....                          | 37        |
| 5.4.          | ENTRENAMIENTO DE LA RED .....  | 39        |
| 5.4.1.        | INFLUENCIA DEL TAMAÑO DE LOTE EN LA EFICIENCIA DE EFFICIENTNETV2 ..... | 40        |
| 5.4.2.        | COMPARACIÓN EFFICIENTNETV2, RESNET50, VGG16 .....                      | 44        |
| 5.4.3.        | ANÁLISIS DEL TIEMPO DE ENTRENAMIENTO .....                             | 51        |
| 5.5.          | VALIDACIÓN DEL MODELO CON CONJUNTO DE PRUEBA INDEPENDIENTE.....        | 52        |
| 5.6.          | DOCUMENTACIÓN DE LOS RESULTADOS OBTENIDOS .....                        | 53        |
| 5.6.1.        | VALIDACIÓN MEDICA .....  | 63        |
| 5.7.          | REGISTRO SOBRE LA EFECTIVIDAD DEL MODELO .....                         | 65        |
| 5.8.          | LIMITACIONES .....   | 67        |
| 5.9.          | ÁREAS DE MEJORA .....  | 67        |
| <b>VI.</b>    | <b>CONCLUSIONES</b> .....  | <b>69</b> |
| <b>VII.</b>   | <b>REFERENCIAS BIBLIOGRÁFICAS</b> .....                                | <b>71</b> |
| <b>ANEXOS</b> | .....  | <b>76</b> |

## ÍNDICE DE ILUSTRACIONES

|                        |   |           |
|------------------------|---|-----------|
| <b>Ilustración 1.</b>  | Línea de Tiempo .....                                   | <b>6</b>  |
| <b>Ilustración 2.</b>  | Palabras claves predominantes en los antecedentes ..... | <b>10</b> |
| <b>Ilustración 3.</b>  | Distribución global de estudios.....                    | <b>16</b> |
| <b>Ilustración 4.</b>  | Imagen Integradora de Proyecto .....                    | <b>18</b> |
| <b>Ilustración 5.</b>  | Variables de Investigación .....                        | <b>28</b> |
| <b>Ilustración 6.</b>  | Metodología en V .....                                  | <b>30</b> |
| <b>Ilustración 7.</b>  | Resumen del procedimiento.....                          | <b>33</b> |
| <b>Ilustración 8.</b>  | Precisión lote 8 Efficienetv2 .....                     | <b>41</b> |
| <b>Ilustración 9.</b>  | Perdida lote 8 Efficientnetv2.....                      | <b>41</b> |
| <b>Ilustración 10.</b> | Precisión lote 12 Efficienetv2.....                     | <b>42</b> |
| <b>Ilustración 11.</b> | Perdida lote 12 Efficientnetv2 .....                    | <b>42</b> |
| <b>Ilustración 12.</b> | Precisión lote 32 Efficienetv2.....                     | <b>43</b> |
| <b>Ilustración 13.</b> | Perdida lote 32 Efficientnetv2 .....                    | <b>43</b> |
| <b>Ilustración 14.</b> | Precisión Efficienetv2.....                             | <b>46</b> |
| <b>Ilustración 15.</b> | Perdida Efficientnetv2.....                             | <b>46</b> |
| <b>Ilustración 16.</b> | Precisión ResNet50v2 .....                              | <b>48</b> |
| <b>Ilustración 17.</b> | Perdida ResNet50v2 .....                                | <b>48</b> |
| <b>Ilustración 18.</b> | Precisión VGG16.....                                    | <b>49</b> |
| <b>Ilustración 19.</b> | Perdida VGG16 .....                                     | <b>50</b> |

|   |    |
|---|----|
| <b>Ilustración 20.</b> Curva ROC EfficientNetv2 .....                                 | 55 |
| <b>Ilustración 21.</b> Curva ROC ResNet50v2.....                                      | 55 |
| <b>Ilustración 22.</b> Matriz de Confusión EfficientNetv2.....                        | 56 |
| <b>Ilustración 23.</b> Matriz de Confusión ResNetv2.....                              | 56 |
| <b>Ilustración 24.</b> Matriz de Confusión Normalizada EfficientNetv2 .....           | 57 |
| <b>Ilustración 25.</b> Matriz de Confusión Normalizada ResNet50v2.....                | 58 |
| <b>Ilustración 26.</b> Distribución de errores y predicciones EfficientNetv2.....     | 59 |
| <b>Ilustración 27.</b> Distribución de errores y predicciones ResNet50v2 .....        | 59 |
| <b>Ilustración 28.</b> Curva Precision-Recall EfficientNetV2.....                     | 60 |
| <b>Ilustración 29.</b> Curva Precision-Recall ResNet50v2.....                         | 61 |
| <b>Ilustración 30.</b> Distribución de predicciones ResNet50v2 y EfficientNetv2 ..... | 63 |
| <b>Ilustración 31.</b> Interfaz del Prototipo .....                                   | 64 |
| <b>Ilustración 32.</b> Interfaz del Prototipo .....                                   | 65 |
| <b>Ilustración 33.</b> Épocas tamaño de lote 8 EfficientNetv2.....                    | 76 |
| <b>Ilustración 33.</b> Épocas tamaño de lote 6 EfficientNetv2.....                    | 76 |
| <b>Ilustración 35.</b> Épocas tamaño de lote 12 EfficientNetv2.....                   | 77 |
| <b>Ilustración 36.</b> Épocas tamaño de lote 32 EfficientNetv2.....                   | 77 |
| <b>Ilustración 37.</b> Épocas tamaño de lote 6 ResNet50v2 .....                       | 77 |
| <b>Ilustración 38.</b> Arquitectura ResNet50v2.....                                   | 77 |
| <b>Ilustración 39.</b> Arquitectura EfficientNetv2 .....                              | 78 |

## ÍNDICE DE TABLAS

|  |    |
|--|----|
| <b>Tabla 1.</b> Limitaciones otros investigadores .....                                      | 19 |
| <b>Tabla 2.</b> Objetivo del Método .....  | 26 |
| <b>Tabla 3.</b> Objetivo del Método .....  | 26 |
| <b>Tabla 4.</b> Cronograma de Actividades .....  | 35 |
| <b>Tabla 5.</b> Distribución de Datos.....   | 38 |
| <b>Tabla 6.</b> Resultados de tamaños de lotes.....  | 44 |
| <b>Tabla 7.</b> Resultados de modelos.....   | 50 |
| <b>Tabla 8.</b> Resultados de modelos.....   | 52 |
| <b>Tabla 7.</b> Métricas de Clasificación EfficientNetV2 (16 épocas, tamaño de lote 6) ..... | 54 |
| <b>Tabla 8.</b> Métricas de Clasificación ResNet50v2 (16 épocas, tamaño de lote 6) .....     | 60 |
| <b>Tabla 11.</b> Comparación de Resultados con Antecedentes.....                             | 65 |

## ÍNDICE DE ECUACIONES

|   |    |
|---|----|
| <b>Ecuación 1.</b> Fórmula para el cálculo de precisión .....       | 34 |
| <b>Ecuación 2.</b> Fórmula para el cálculo de la sensibilidad ..... | 34 |
| <b>Ecuación 3.</b> Fórmula para el cálculo del f1-score.....        | 34 |

## LISTA DE SIGLAS

|     |                                   |
|-----|-----------------------------------|
| BC  | Breast Cancer                     |
| TN  | True Negatives                    |
| CNN | Convolutional Neural Network      |
| API | Application Programming Interface |
| GPU | Graphics Processing Unit          |
| CPU | Central Processing Unit           |
| IA  | Inteligencia Artificial           |
| IOT | Internet of Things                |
| TP  | True Positives                    |
| DBT | Digital Breast Tomosynthesis      |
| FP  | False Positives                   |
| CAD | Computer-Aided Diagnosis          |
| ROC | Receiver Operating Characteristic |
| AUC | Area Under the Curve              |

## GLOSARIO

1. Cáncer de Mama: Cáncer que se forma en los tejidos de la mama.(NIH, 2011)
2. Inteligencia Artificial (IA): Tecnología que permite que las computadoras simulen la inteligencia humana y las capacidades humanas de resolución de problema.(IBM, 2021)
3. Red Neuronal Convolucional : Es una arquitectura de red para Deep Learning que aprende directamente a partir de datos.(MATLAB & Simulink, 2022)
4. Ecografía Mamaria: Técnica de diagnóstico que usa los ultrasonidos para formar una imagen de la estructura de las mamas.(Universidad de Navarra, 2023)
5. Sobreajuste : Ocurre cuando un algoritmo se ajusta demasiado o incluso exactamente a sus datos de entrenamiento, lo que da como resultado un modelo que no puede hacer predicciones o conclusiones precisas a partir de ningún otro dato que no sea el de entrenamiento.(IBM, 2024)
6. Transfer Learning: enfoque de Deep Learning que utiliza un modelo entrenado para una tarea como punto de partida para otro modelo que realiza una tarea similar.(MATLAB & Simulink,2022b)
7. Tumor: Masa de células transformadas, con crecimiento y multiplicación anormales(RAE, 2024)
8. Tumor benigno: Tumor no canceroso, formado por células muy semejantes a las normales y que no se extiende a otros tejidos u órganos(RAE, 2024b)
9. Tumor maligno: Tumor canceroso y que se extiende a otros tejidos u órganos.(RAE, 2024c)

## I. INTRODUCCIÓN

El cáncer de mama es una de las principales causas de muerte entre las mujeres en todo el mundo (Cáncer de Mama, 2024). A pesar de los avances en el diagnóstico por imagen, muchos casos se detectan en etapas avanzadas, cuando el tratamiento es menos efectivo y más costoso (Naseem et al., 2022). La interpretación de las imágenes mamográficas en general requiere de personal especializado y puede estar sujeta a errores humanos. En este contexto, la inteligencia artificial (IA) surge como una herramienta para asistir y mejorar la precisión diagnóstica en la interpretación de imágenes (Dileep & Gianchandani Gyani, 2022).

Este proyecto se centrará en evaluar los modelos de inteligencia artificial existentes para la clasificación de tumores en imágenes de ecografías mamarias. A través de la evaluación comparativa de diferentes arquitecturas de redes neuronales convolucionales y la implementación de técnicas de preprocesamiento y validación, se buscará mejorar la precisión en la identificación de tumores. Se seleccionará un modelo de IA robusto que pueda clasificar con precisión los tipos de cáncer de mama y proporcionar una herramienta de apoyo eficaz para los profesionales de la salud.

El estudio se estructura en varias fases clave: la selección y preprocesamiento de una base de datos diversificada de imágenes de ecografías mamarias, el análisis de la influencia del tamaño del lote en el entrenamiento del modelo, la comparación de diferentes arquitecturas de redes neuronales y la validación del modelo final utilizando un conjunto de datos independiente. Este enfoque sistemático y tecnológicamente subraya el valor de la IA en la mejora de los diagnósticos, ofreciendo en la identificación temprana del cáncer de mama.

Finalmente, se validará el modelo final utilizando un conjunto de datos independiente. Esta validación permitirá evaluar la capacidad de generalización y precisión del modelo en la práctica clínica, asegurando que el modelo pueda proporcionar diagnósticos precisos y confiables en un entorno real.

El estudio se compondrá de 8 capítulos de cada fase específica del proyecto. El capítulo II, "Estado del Arte," describirá los antecedentes y los estudios previos relacionados con el uso de IA en la detección del cáncer de mama. El capítulo III, "Objetivos," presentará los objetivos específicos

del estudio, detallando lo que se pretende lograr con esta investigación. El capítulo IV, "Métodos," explicará los materiales y métodos utilizados, incluyendo la selección de datos y las técnicas de preprocesamiento. El capítulo V, "Resultados," discutirá la configuración experimental, los principales resultados obtenidos y su análisis. Finalmente, el capítulo VI, "Conclusiones," ofrecerá un resumen de los hallazgos del estudio, destacando las implicaciones de los mismos y las posibles direcciones futuras para la investigación en este campo.

## **II. ESTADO DEL ARTE**

En este capítulo, se revisarán los avances y las investigaciones previas en el uso de la inteligencia artificial (IA) para la detección del cáncer de mama en imágenes de ecografías mamarias. Se analizarán diversos enfoques y técnicas empleadas en estudios anteriores, incluyendo el uso de redes neuronales convolucionales (CNN) y el aprendizaje por transferencia. Se explorará cómo estas metodologías han contribuido a mejorar la precisión diagnóstica y se evaluarán las limitaciones y desafíos identificados en la literatura.

### **2.1. ANTECEDENTES**

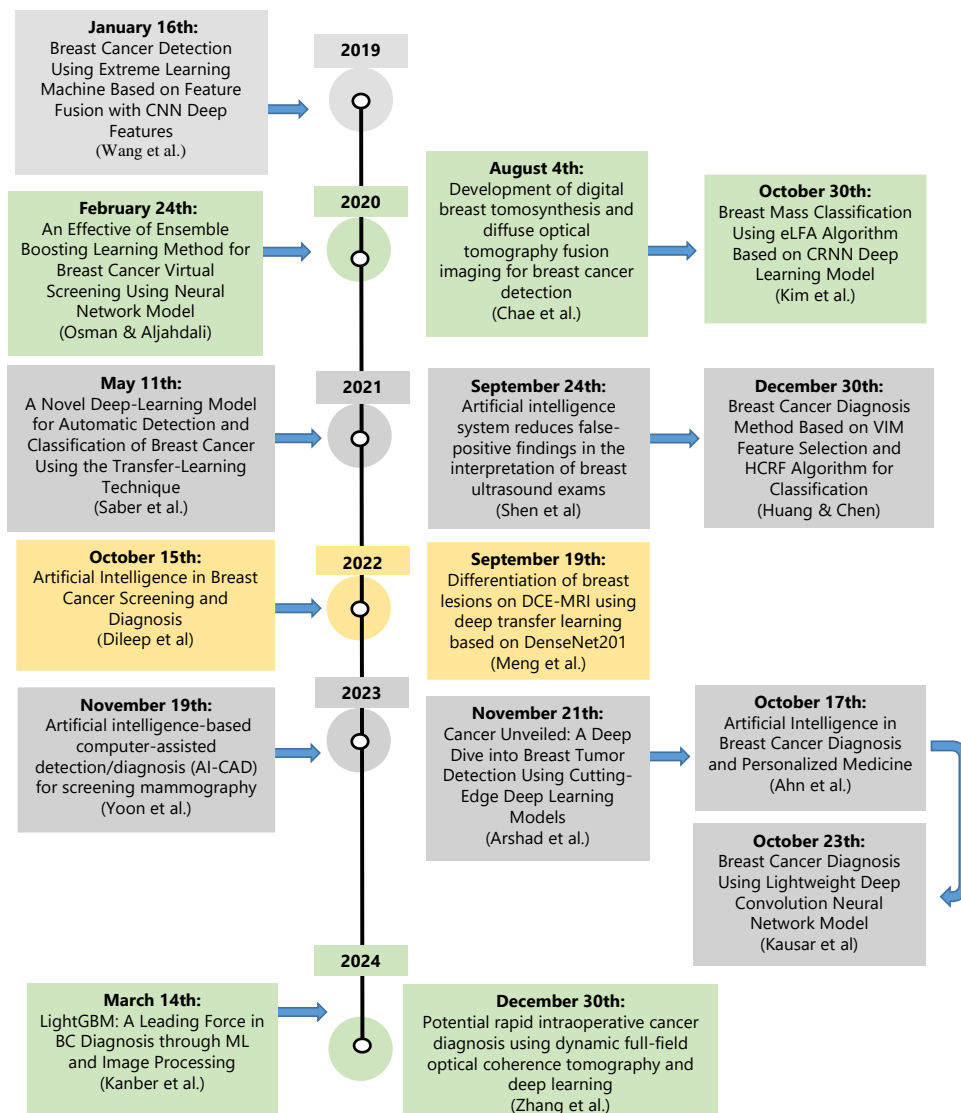
La indagación de la literatura académica reveló trabajos investigativos relacionados a la temática del presente trabajo en el periodo de 2019-2024 (Ilustración 1). La revisión sistemática de estos estudios demuestra un creciente interés en la integración de tecnologías emergentes en el diagnóstico y tratamiento de enfermedades, con un énfasis particular en la inteligencia artificial y su aplicación en el campo de la salud. A lo largo de este periodo, se han identificado avances significativos en la precisión de los algoritmos de aprendizaje profundo, la interpretación de grandes conjuntos de datos biomédicos y la personalización de terapias basadas en perfiles genéticos y moleculares.

En el ámbito de la aplicación de inteligencia artificial para el diagnóstico de cáncer de mama, se destacó el estudio de Wang et al., (2019), «Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion With CNN Deep Features», abordó un enfoque experimental en la detección del cáncer de mama, utilizando técnicas de inteligencia artificial para mejorar la identificación de masas mamarias. El trabajo se centró en la fusión de características y la implementación de modelos de Extreme Learning Machine (ELM) para entrenar sistemas de inteligencia artificial.

El objetivo principal de la investigación fue desarrollar un método innovador que mejorara la detección del cáncer de mama combinando características de ELM y características profundas de redes neuronales convolucionales (CNN). El estudio se enfocó en el desarrollo de métodos de detección de masas y diagnóstico de cáncer de mama, utilizando técnicas avanzadas de

aprendizaje automático y fusión de características para lograr un rendimiento superior en la detección y diagnóstico de la enfermedad.

Las conclusiones del estudio resaltaron la propuesta de un método de diagnóstico asistido por computadora (CAD) para el cáncer de mama basado en la fusión de características profundas, aplicando características extraídas de CNN en las etapas de detección y diagnóstico de masas. Se enfatizó la importancia de la selección de características en la precisión del diagnóstico y la superioridad de la fusión de características profundas en comparación con métodos anteriores.



**Ilustración 1.** Línea de Tiempo

Fuente: Elaboración propia utilizando los títulos de las investigaciones mencionada en los antecedentes.

En el estudio realizado por de Osman & Aljahdali, (2020), titulado «An Effective of Ensemble Boosting Learning Method for Breast Cancer Virtual Screening Using Neural Network Model», fue una investigación experimental que destacó por su enfoque en la mejora de la detección del cáncer de mama a través de la inteligencia artificial. Se empleó el método de Redes Neuronales de Función Radial (RBFNN) para refinar el proceso de predicción y aumentar la precisión del diagnóstico.

El objetivo de la investigación fue profundizar y optimizar el proceso de diagnóstico del cáncer de mama mediante la integración de estas técnicas, buscando alcanzar una mayor exactitud en la detección de la enfermedad.

El estudio logró una tasa de precisión del 98.4% en el conjunto de datos de diagnóstico de cáncer de mama. En el estudio se mencionó como mejora la creación de un sistema de predicción inteligente que permitiera una interacción directa con los pacientes y la enfermedad, sin la necesidad de intervención médica, lo que podría haber facilitado una detección temprana del cáncer y mejorado los resultados clínicos.

El estudio prospectivo realizado por Chae et al., (2020), titulado «Development of digital breast tomosynthesis and diffuse optical tomography fusion imaging for breast cancer detection», representó un avance significativo en la detección del cáncer de mama. La investigación se centró en la aplicación de técnicas de imagen de fusión DBT/DOT, con un proceso metodológico que incluyó la adquisición de imágenes de DB y DOT en 28 mujeres y su posterior análisis independiente por dos radiólogos especializados en mamografías.

El objetivo de la investigación fue evaluar la utilidad de la técnica de imagen de fusión DBT/DOT para la detección del cáncer de mama, buscando obtener datos preliminares sobre su eficacia diagnóstica. El estudio concluyó que la imagen de fusión DBT/DOT mejora significativamente el rendimiento diagnóstico en comparación con el uso independiente de DBT o DOT. Las áreas bajo la curva característica de operación del receptor (AUC) mejoraron notablemente con la técnica de fusión, sugiriendo que esta podría ser una herramienta valiosa en el diagnóstico del cáncer de mama. Además, la imagen de fusión DBT/DOT demostró tener los resultados más altos, lo que indica una mayor la detección de lesiones mamarias malignas.

El estudio «Breast Mass Classification Using eLFA Algorithm Based on CRNN Deep Learning Model» realizado por Kim et al., (2020), es una investigación computacional que aplicó técnicas de inteligencia artificial para la detección de masas mamarias. Se utilizó el modelo eLFA-CRNN y algoritmos de detección de bordes como Canny, Sobel-mask, Laplacian y HOG.

El objetivo fue desarrollar un modelo de aprendizaje profundo basado en el algoritmo eLFA para clasificar masas mamarias en imágenes ultrasónicas, buscando mejorar la precisión en la detección del cáncer de mama.

Las conclusiones del estudio indicaron que el algoritmo eLFA, junto con el modelo eLFA-CRNN, alcanzó una precisión del 98% en la clasificación de masas benignas y malignas, superando a otros algoritmos de detección de bordes. La utilización de segmentos de líneas en la generación de mapas de características fue clave para la precisión del modelo en la detección de masas en imágenes ultrasónicas de senos.

En el estudio del autor Saber et al., (2021), «A Novel Deep-Learning Model for Automatic Detection and Classification of Breast Cancer Using the Transfer-Learning Technique», representa un avance significativo en la detección y clasificación del cáncer de mama mediante el uso de técnicas de inteligencia artificial. La investigación se centró en la implementación de modelos de aprendizaje profundo, en particular la red neuronal convolucional (CNN) VGG16, complementada con estrategias de transferencia de aprendizaje para optimizar la precisión en la identificación de lesiones mamarias.

El propósito del estudio fue desarrollar un modelo de inteligencia artificial que pudiera asistir a los médicos en la detección y diagnóstico del cáncer de mama, utilizando datos del conjunto MIAS y aplicando técnicas avanzadas de procesamiento de imágenes y aprendizaje automático.

Las conclusiones y contribuciones más destacadas del estudio incluyeron la creación de un modelo de aprendizaje profundo que alcanzó altos niveles de precisión, sensibilidad, especificidad, puntaje F y AUC en la detección y clasificación del cáncer de mama. La integración de la CNN a través de la transferencia de aprendizaje resultó ser efectiva para mejorar los resultados en comparación con métodos anteriores, lo que indicó un progreso importante en la

aplicación de la inteligencia artificial en la medicina para la detección temprana y precisa de esta enfermedad.malignas.

El estudio de Shen et al., (2021), «Artificial intelligence system reduces false-positive findings in the interpretation of breast ultrasound exams», se enfocó en el uso de la inteligencia artificial para minimizar los hallazgos falsos positivos en la interpretación de ultrasonidos mamarios. Se entrenaron modelos de IA con un amplio conjunto de datos y se implementó un sistema de aprendizaje débilmente supervisado para clasificar lesiones en imágenes de ultrasonido. El objetivo fue evaluar la eficacia de un sistema de IA en la reducción de falsos positivos en ultrasonidos mamarios, buscando mejorar la precisión en la detección de lesiones malignas y disminuir las biopsias benignas innecesarias.

Las conclusiones del estudio resaltaron la alta precisión diagnóstica del sistema de IA, con un área bajo la curva ROC (AUROC) de 0.976. Además, se destacó la capacidad del sistema para reducir los falsos positivos y las biopsias benignas, y la precisión en la interpretación de ultrasonidos mamarios.

El estudio experimental realizado por Huang & Chen, (2021), titulado «Breast Cancer Diagnosis Method Based on VIM Feature Selection and HCRF Algorithm for Classification», representó un avance significativo en la detección del cáncer de mama mediante la aplicación de técnicas de inteligencia artificial. La investigación se centró en la selección de características a través de la Variable Importance Measure (VIM) y la clasificación utilizando el algoritmo Hierarchical Clustering Random Forest (HCRF).

El objetivo principal fue desarrollar un modelo de diagnóstico que no solo mejorara la precisión y eficiencia en la detección del cáncer de mama, sino que también redujera la complejidad y el tiempo de prueba del modelo.

Las conclusiones del estudio alcanzaron la precisión de 97.05% en el conjunto de datos WDBC y el 97.76% en el conjunto de datos WBC. La combinación de VIM y HCRF resultó en un rendimiento mejorado y una mayor capacidad de generalización del clasificador, superando a los modelos de random forest tradicionales. Estos hallazgos resultaron favorables para la detección temprana del cáncer de mama, lo que podría tener un impacto directo en salvar vidas mediante

diagnósticos más oportunos y precisos. Además, el método propuesto mostró un avance para la detección de otros tipos de cáncer y podría guiar a los médicos en el diagnóstico precoz de tumores mamarios.

Las palabras clave predominantes derivadas de la investigación académica sobre el diagnóstico y tratamiento del cáncer de mama en el período de 2019 a 2024 reflejaron las áreas de enfoque más relevantes y los avances significativos en este campo. Estas indicaron las prioridades y tendencias emergentes en la investigación del cáncer de mama. A través del análisis de estas palabras, se observó un énfasis en aspectos cruciales como la detección temprana y el uso de la inteligencia artificial. (Ilustración 2).



**Ilustración 2.** Palabras claves predominantes en los antecedentes

Fuente: Elaboración propia

Siguiendo la línea de avance científico el estudio Dileep et al., (2022) en su artículo «Artificial Intelligence in Breast Cancer Screening and Diagnosis» Dicha investigación tecnológica abordó el uso de técnicas de inteligencia artificial para la detección de tumores mamarios, empleando modelos de inteligencia artificial basados en machine learning y deep learning para analizar imágenes médicas y detectar la presencia de tumores en pacientes con cáncer de mama.

Se recopilaron datos de estudios de casos previos y se entrenaron los algoritmos con un amplio conjunto de imágenes mamográficas para mejorar la precisión en la identificación de lesiones malignas, utilizando técnicas de segmentación y clasificación de tumores para diferenciar entre tejido sano y maligno con alta exactitud.

El objetivo general de dicha investigación fue evaluar la eficacia de los algoritmos de inteligencia artificial en la identificación precisa de la presencia tumoral en pacientes con cáncer de mama, con el propósito de mejorar la precisión y rapidez en el diagnóstico de esta enfermedad.

Entre las principales conclusiones del estudio se encontró la capacidad de la inteligencia artificial para detectar de manera temprana la presencia de tumores mamarios, lo que podría contribuir significativamente a la mejora de los resultados clínicos y la supervivencia de las pacientes con cáncer de mama.

La investigación llevada a cabo por Meng et al., (2022), titulada «Differentiation of breast lesions on dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) using deep transfer learning based on DenseNet201», constituyó un estudio retrospectivo que incursionó en el campo del aprendizaje profundo para mejorar el diagnóstico de cáncer de mama. El estudio se valió de técnicas de Deep Transfer Learning (DTL), específicamente del modelo DenseNet201, para diferenciar lesiones mamarias en imágenes de resonancia magnética dinámica con contraste. La implementación de la inteligencia artificial se realizó utilizando Python, Keras y TensorFlow, en un entorno de desarrollo equipado con Windows 10, procesador Intel Core i7-10700F y tarjeta gráfica NVIDIA GeForce GTX 2060.

El propósito central de la investigación fue evaluar la eficacia de diversas estrategias de ajuste fino de DTL basadas en DenseNet201, con el fin de distinguir entre lesiones benignas y malignas en imágenes de DCE-MRI de mama.

Las conclusiones del estudio resaltaron las estrategias de ajuste fino de DenseNet201 para la diferenciación precisa de lesiones mamarias. Estas estrategias mostraron una mejora de la precisión de clasificación de lesiones, sugiriendo su aplicabilidad en la práctica clínica para el diagnóstico de cáncer de mama. El estudio subrayó la importancia de la inteligencia artificial y el

aprendizaje profundo en el avance del diagnóstico médico, ofreciendo nuevas perspectivas para la detección y clasificación de enfermedades en imágenes médicas.

El estudio realizado por Yoon et al., (2023), titulado «Artificial intelligence-based computer-assisted detection/diagnosis (AI-CAD) for screening mammography: Outcomes of AI-CAD», se empleó un enfoque retrospectivo y transversal. La investigación implicó el uso de un software AI-CAD disponible comercialmente para analizar mamografías, revisar lecturas históricas de tres radiólogos para su interpretación y evaluar retrospectivamente las características anormales detectadas por AI-CAD. La verdad fundamental sobre el cáncer, los hallazgos benignos o la ausencia de anomalías se confirmó mediante diagnóstico histopatológico o resultados negativos en exámenes de detección posteriores.

El principal objetivo de la investigación fue evaluar el rendimiento diagnóstico independiente de AI-CAD y los resultados de las anomalías detectadas por AI-CAD dentro del flujo de trabajo de interpretación mamográfica. El estudio de AI-CAD, detectó que el 17,9% de los cánceres adicionales inicialmente pasados por alto, por los radiólogos en una población de detección consecutiva. Además, AI-CAD mostró tasas de recuperación significativamente más altas en comparación con las interpretaciones de los radiólogos, y el 89 % de las calificaciones de AI-CAD resultaron negativas. Las diferencias en las características anormales detectadas por los lectores humanos y AI-CAD resaltan la importancia de seguir investigando cómo estas disparidades afectan la interpretación para la posible implementación de AI-CAD en el flujo de trabajo de interpretación.

El estudio llevado a cabo Arshad et al., (2023), titulado «Cancer Unveiled: A Deep Dive Into Breast Tumor Detection Using Cutting-Edge Deep Learning Models», fue una investigación de revisión que se centró en la aplicación de modelos de transferencia de aprendizaje en inteligencia artificial para la detección de carcinoma ductal invasivo (IDC) en imágenes histopatológicas. Se destacaron en este estudio el uso de arquitecturas de aprendizaje profundo como VGG-16, DenseNet-121 y MobileNetV2, y la evaluación de su precisión de clasificación.

El propósito de la investigación fue realizar una comparativa de la efectividad de distintos modelos de inteligencia artificial, con especial atención en la detección de IDC a partir de

imágenes histopatológicas, enfocándose en la aplicación de técnicas de transferencia de aprendizaje.

Las conclusiones del estudio resaltaron una alta precisión de clasificación alcanzada por los modelos VGG-16 y DenseNet-121, con un 98% y un 99% respectivamente, lo que demostró un diagnóstico preciso y eficiente de IDC. Además, se enfatizó la utilidad de la transferencia de aprendizaje en el análisis de imágenes médicas y se sugirió la exploración de una gama más amplia de arquitecturas en futuras investigaciones para continuar avanzando en el aprendizaje profundo aplicado a la detección del cáncer de mama.

El estudio de revisión realizado Ahn et al., (2023), «Artificial Intelligence in Breast Cancer Diagnosis and Personalized Medicine», abordó la aplicación de la inteligencia artificial (IA) en el diagnóstico del cáncer de mama y la medicina personalizada. Se centró en el análisis de imágenes médicas como mamografías, resonancias magnéticas y ultrasonidos, utilizando algoritmos de aprendizaje profundo, como las redes neuronales convolucionales, para identificar patrones y características relevantes.

El objetivo de la investigación fue investigar cómo la IA puede transformar el diagnóstico del cáncer de mama, mejorando la precisión del diagnóstico, la evaluación de biomarcadores, la predicción de la respuesta al tratamiento y la personalización de las estrategias terapéuticas para cada paciente.

Las conclusiones del estudio destacaron la capacidad de la IA para revolucionar el diagnóstico del cáncer de mama, permitiendo una detección más temprana y precisa, y facilitando la personalización de los tratamientos basados en las características individuales de cada paciente. Además, se resaltó la capacidad de la IA para mejorar la atención médica en oncología, ofreciendo nuevas posibilidades para la medicina personalizada y la mejora de los resultados clínicos en pacientes con cáncer de mama.

El estudio de Kausar et al., (2023), «Breast Cancer Diagnosis Using Lightweight Deep Convolution Neural Network Model», es una investigación experimental que se enfoca en la detección del cáncer de mama mediante el uso de inteligencia artificial. Se destacan las técnicas de transformada wavelet para el muestreo descendente de imágenes histopatológicas de alta

resolución y la implementación de un modelo de CNN ligero para la extracción de características profundas.

El objetivo de la investigación fue desarrollar un método eficiente y preciso para la clasificación del cáncer de mama a partir de imágenes histopatológicas, superando los desafíos computacionales de los modelos de deep learning en aplicaciones en tiempo real.

Las conclusiones del estudio indicaron una reducción del costo computacional gracias a la transformada wavelet y al modelo de CNN ligero, manteniendo la precisión en la clasificación de diferentes clases de cáncer. La robustez del enfoque propuesto, en comparación con técnicas convencionales, sugiere que la detección automática de cáncer de mama y su aplicación en otras áreas de patología computacional.

El estudio realizado por Kanber et al., (2024), titulado «LightGBM: A Leading Force in BC Diagnosis Through ML and Image Processing», fue una investigación exploratoria llevada a cabo en 2024 que destacó por su enfoque innovador en el diagnóstico del cáncer de mama. La investigación se centró en el análisis de histogramas de color, la extracción de características de contorno y momentos invariantes en imágenes histopatológicas del conjunto de datos BreakHis, utilizando una variedad de modelos de inteligencia artificial, incluyendo LightGBM, CatBoost, XGBoost, NaiveBayes, DT, SVM, RF, QDA, GPC y regresión de ridge.

El objetivo principal fue desarrollar un flujo de trabajo robusto para la extracción de características que integrara diversas técnicas, generando un vector de características unificado que permitiera una representación detallada de los tejidos. Se evaluaron múltiples algoritmos de aprendizaje automático en diferentes ampliaciones de imágenes, enfocándose en su rendimiento diagnóstico en varios escenarios clínicos.

Las conclusiones del estudio resaltaron la superioridad de LightGBM en términos de eficiencia computacional y precisión diagnóstica, demostrando que el aumento de datos mejora la capacidad de clasificación de los modelos. Demostró fortalecer la robustez y generalización de los modelos de aprendizaje automático en el análisis de imágenes médicas. El estudio subrayó la importancia de equilibrar el rendimiento con las demandas computacionales al seleccionar

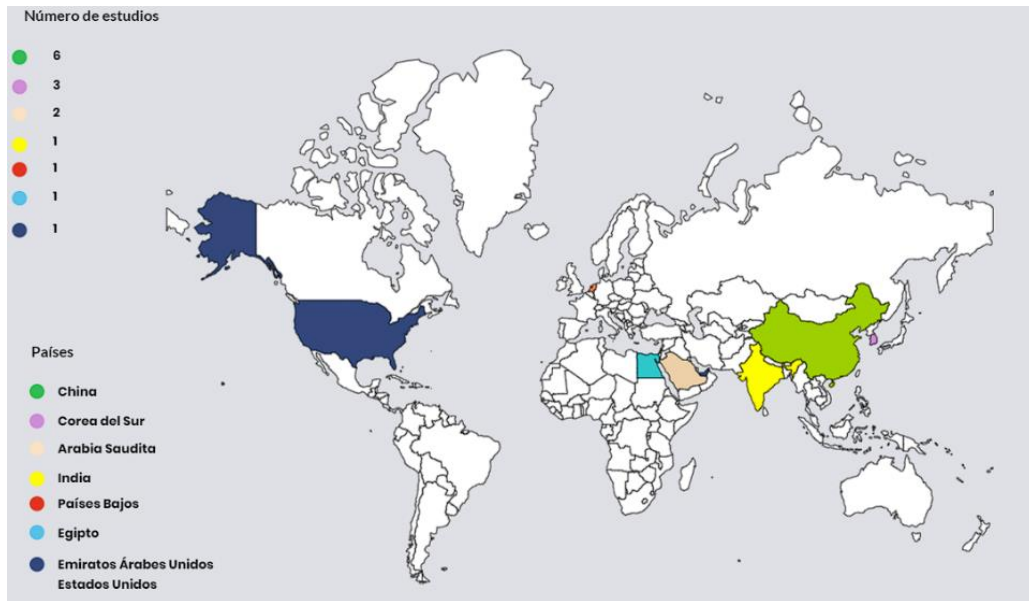
algoritmos para aplicaciones clínicas y destacó el impacto positivo de los enfoques mejorados de aprendizaje automático en el diagnóstico de cáncer de mama.

El estudio prospectivo realizado por Zhang et al., (2024), titulado «Potential rapid intraoperative cancer diagnosis using dynamic full-field optical coherence tomography and deep learning: A prospective cohort study in breast cancer patients», representó un avance significativo en el campo del diagnóstico intraoperatorio del cáncer de mama. La investigación combinó la tomografía de coherencia óptica de campo completo dinámica (D-FFOCT) con el aprendizaje profundo para desarrollar un modelo de inteligencia artificial capaz de diagnosticar el cáncer de mama durante la cirugía. Se utilizaron 10,357 parches de D-FFOCT para entrenar el modelo, lo que subraya la amplitud de datos empleados para asegurar la precisión del sistema.

El objetivo fue crear un flujo de trabajo automatizado y no destructivo que mejorara la eficiencia y precisión de los procedimientos de diagnóstico durante la cirugía. El estudio logró una tasa de precisión del 97.62% en el diagnóstico automatizado del cáncer de mama, y la simulación del diagnóstico de márgenes intraoperatorios se completó en solo tres (3) minutos. Estos resultados son notables, ya que indican una mejora considerable en comparación con los métodos tradicionales de diagnóstico.

La investigación demostró que la combinación de D-FFOCT y aprendizaje profundo podría revolucionar el diagnóstico intraoperatorio del cáncer, ofreciendo una alternativa más rápida y precisa a los métodos convencionales de laboratorio de patología.

Se consultaron diferentes artículos sobre la identificación de presencia tumoral en pacientes con cáncer de mama mediante la aplicación de la inteligencia artificial. La indagación literaria efectuada incluyó países como China, Corea del Sur, Arabia Saudita, India, Países Bajos, Egipto, Emiratos Árabes Unidos y Estados Unidos (Ilustración 3).



**Ilustración 3.** Distribución global de estudios

Fuente: Elaboración propia con fundamento en las regiones de los artículos consultados

## 2.2. PROBLEMÁTICA

El cáncer de mama es el segundo tipo de cáncer más prevalente entre las mujeres y conduce a la muerte, lo que representa un desafío significativo en la salud pública mundial (Naseem et al., 2022). Según datos de la Agencia Internacional para la Investigación del Cáncer (IARC) indicaron que el cáncer de mama ha superado al de pulmón como el cáncer más diagnosticado en todo el mundo, con una estimación de 2,26 millones de nuevos casos y aproximadamente 685.000 muertes sólo en 2020. (Kanber et al., 2024).

El cáncer de mama es más eficaz de tratar si se diagnostica a tiempo. La eficacia del tratamiento en etapas posteriores es pobre. Por lo tanto, el diagnóstico temprano y la prevención pueden ser útiles para registrar más casos y reducir el número de muertes (Dileep & Gianchandani Gyani, 2022). El cáncer de mama es una de las principales causas de muerte entre las mujeres en todo el mundo. A pesar de los avances en el diagnóstico por imagen, muchos casos se detectan en etapas avanzadas, cuando el tratamiento es menos efectivo y más costoso. La falta de recursos, como patólogos competentes, provoca un retraso en el diagnóstico y una planificación inadecuada de la terapia, todo lo cual contribuye a una tasa de supervivencia pésima (Arshad et al., 2023). Esto destaca un problema crítico en el diagnóstico y tratamiento del cáncer de mama:

la escasez de recursos, particularmente la falta de patólogos competentes. Esta deficiencia puede llevar a retrasos significativos en el diagnóstico, lo que a su vez puede afectar negativamente la planificación de la terapia. Estos retrasos y la planificación inadecuada pueden contribuir a una disminución en las tasas de supervivencia de los pacientes con cáncer de mama.

Para profundizar en esta problemática, es importante considerar el impacto que tiene la falta de recursos en el sistema de salud y en los resultados de los pacientes. Según la Organización Mundial de la Salud, el retraso o la inaccesibilidad de la atención del cáncer conduce a una menor probabilidad de supervivencia, mayor morbilidad del tratamiento y mayores costos de atención, lo que resulta en muertes y discapacidades evitables por cáncer (OMS, 2022.). Además, estudios han demostrado que incluso un retraso de cuatro semanas en el tratamiento del cáncer puede aumentar el riesgo de mortalidad (Hanna et al., 2020).

La interpretación de las imágenes mamográficas requiere de personal especializado y puede estar sujeta a errores humanos. Por ello, existe la necesidad de desarrollar sistemas automatizados que puedan ayudar a identificar y segmentar los tumores de forma rápida y precisa, utilizando técnicas de inteligencia artificial. Estos sistemas podrían ayudar a reducir la carga de trabajo de los radiólogos, aumentar la consistencia en la interpretación de las imágenes y, lo más importante, mejorar la precisión en la identificación de lesiones malignas. La IA tiene el potencial de analizar grandes volúmenes de datos rápidamente y con alta precisión, lo que podría ser crucial para la detección temprana del cáncer de mama y, por ende, para mejorar las tasas de supervivencia y la calidad de vida de las pacientes (Freeman et al., 2021).

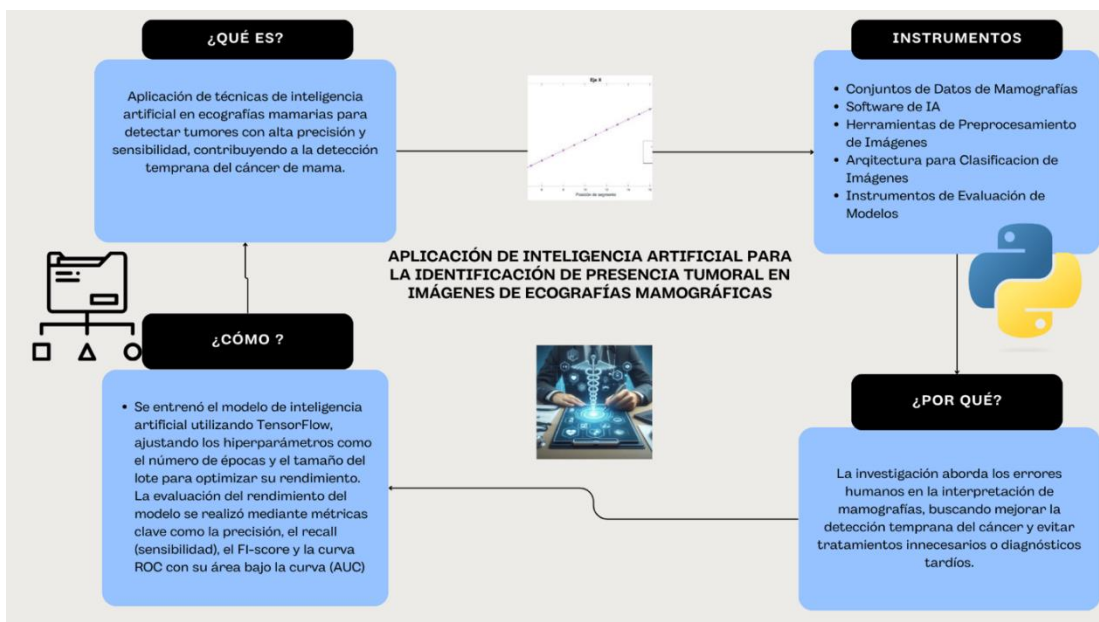
La detección temprana es crucial para mejorar las tasas de supervivencia y la calidad de vida de las pacientes. La mamografía es actualmente la modalidad de imagen más utilizada para la detección del cáncer de mama; ensayos de detección aleatorios anteriores han demostrado su contribución a la reducción de la mortalidad relacionada con el cáncer de mama (Yoon et al., 2023). Los errores de diagnóstico ya sean falsos positivos o falsos negativos, pueden llevar a tratamientos innecesarios o a la falta de detección oportuna de la enfermedad.

Los falsos positivos en mamografías pueden conducir a biopsias innecesarias, ansiedad adicional para las pacientes y un aumento en los costos de atención médica. Por otro lado, los

falsos negativos son aún más preocupantes, ya que pueden retrasar el diagnóstico y tratamiento del cáncer de mama, lo que potencialmente disminuye las tasas de supervivencia (Mao et al., 2024). En este contexto, la inteligencia artificial (IA) surge como una herramienta prometedora para asistir y mejorar la precisión diagnóstica en la interpretación de imágenes mamográficas.

### 2.3. IMAGEN INTEGRADORA

La investigación por desarrollar demuestra la aplicación de la inteligencia artificial (IA) en el diagnóstico del cáncer de mama, destacando la metodología y las herramientas utilizadas en este proceso avanzado. Se detalla la implementación de técnicas de segmentación en ecografías mamarias y algoritmos de clasificación, que son fundamentales para detectar lesiones tumorales con una precisión mejorada, incrementando la sensibilidad y minimizando los falsos positivos. El diagrama desglosa el flujo de trabajo necesario para poder llevar a cabo el modelo de aprendizaje para la detección de la presencia tumoral del cáncer de mama (Ilustración 4). Este enfoque sistemático y tecnológicamente enriquecido subraya el valor de la IA en la mejora de los diagnósticos médicos, ofreciendo un cambio paradigmático en la detección y el tratamiento del cáncer de mama.



**Ilustración 4.** Imagen Integradora de Proyecto

Fuente: Elaboración propia

## 2.4. TABLA DE LIMITACIONES

**Tabla 1.** Limitaciones otros investigadores

| Ítem | Título del Artículo   | Limitación                                     | Definición de la Limitación  | Fuente                  |
|------|---|--|--|-------------------------|
| 1    | Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion With CNN Deep Features                         | Deficiencias en las características utilizadas | Características subjetivas y objetivas con limitaciones en representación y generalización.  | Wang et al., 2019       |
| 2    | An Effective of Ensemble Boosting Learning Method for Breast Cancer Virtual Screening Using Neural Network Model              | Intervención médica requerida                  | Necesidad de datos de laboratorio y apoyo médico para el diagnóstico, lo que puede retrasar la detección.  | Osman & Aljahdali, 2020 |
| 3    | Development of digital breast tomosynthesis and diffuse optical tomography fusion imaging for breast cancer detection         | Tamaño de la muestra                           | El estudio incluyó solamente a 27 participantes, limitando el uso de métodos estadísticos potentes y la interpretación de resultados preliminares. | Chae et al., 2020       |
| 4    | A Novel Deep-Learning Model for Automatic Detection and Classification of Breast Cancer Using the Transfer-Learning Technique | Generalización a diferentes poblaciones        | Modelos AI entrenados en conjuntos de datos específicos pueden tener limitaciones para generalizar a diversas poblaciones.                         | Saber et al., 2021      |

**Tabla 1. (Continuación)**

| <b>Ítem</b> | <b>Título del Artículo</b>  | <b>Limitación</b>                            | <b>Definición de la Limitación</b>   | <b>Fuente</b>         |
|-------------|---|--|--|-----------------------|
| 5           | A Breast Cancer Diagnosis Method Based on VIM Feature Selection and Hierarchical Clustering Random Forest Algorithm                                   | Evaluación en conjuntos de datos específicos | Evaluación y validación del modelo solo en conjuntos de datos WDBC y WBC, lo que podría afectar la generalización.             | Huang & Chen, 2022    |
| 6           | Artificial Intelligence in Breast Cancer Screening and Diagnosis  | Ética y Confidencialidad de los Datos        | Riesgos éticos como la confidencialidad de los datos, violación de la privacidad, autonomía de los pacientes y consentimiento. | Dileep & Gyani (2022) |
|             |   | Tamaño de la muestra                         | Tamaño pequeño de la muestra que afecta la robustez del modelo.  |                       |
|             |   | Falta de diversidad en las lesiones          | Limita la capacidad del modelo para generalizar entre diferentes tipos de lesiones.  |                       |
| 7           | Differentiation of breast lesions on dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) using deep transfer learning based on DenseNet201 | Limitaciones en la representación de datos   | Basado solo en imágenes DCE-MRI, sin otros datos clínicos o modalidades de imagen.   | Meng et al., 2022     |
|             |   | Necesidad de ajuste fino                     | Importancia del ajuste fino para mejorar la precisión en la clasificación de lesiones.   |                       |

**Tabla 1. (Continuación)**

| Ítem | Título del Artículo  | Limitación                              | Definición de la Limitación   | Fuente              |
|------|--|---|---|---------------------|
| 8    | Artificial intelligence-based computer-assisted detection/diagnosis (AI-CAD) for screening mammography | Tamaño de la muestra                    | Se reconoce que el tamaño de la muestra podría considerarse limitado comparando con la diversidad de casos de la práctica clínica habitual. | Yoon et al. (2023)  |
| 9    | Cancer Unveiled: A Deep Dive Into Breast Tumor Detection Using Cutting-Edge Deep Learning Mode         | Tamaño sustancial del conjunto de datos | Uso de un conjunto de datos considerable, pero se beneficiaría de conjuntos más grandes.  | Arshad et al., 2023 |
|      |  | Exploración limitada de arquitecturas   | Foco en tres arquitecturas de aprendizaje profundo, con necesidad de explorar más.  |                     |
|      |  | Necesidad de investigaciones futuras    | Sugerencia de explorar más arquitecturas y conjuntos de datos para mejorar la generalización.   |                     |

**Tabla 1. (Continuación)**

| <b>Ítem</b> | <b>Título del Artículo</b>   | <b>Limitación</b>  | <b>Definición de la Limitación</b>  | <b>Fuente</b>       |
|-------------|--|--|---|---------------------|
| 10          | Artificial Intelligence in Breast Cancer Diagnosis and Personalized Medicine                       | Validación clínica   | La integración de modelos de IA en entornos clínicos del mundo real presenta desafíos más allá de la validación, como la utilidad y la usabilidad. Los modelos de IA deben someterse a una validación rigurosa mediante variedad de criterios de valoración clínicos. | Ahn et al., 2023    |
| 11          | Breast Cancer Diagnosis Using Lightweight Deep Convolution Neural Network Model                    | Pérdida de información durante la descomposición de imágenes<br>Costo computacional de modelos profundos | Pérdida de información en la transformada wavelet que afecta la precisión de los modelos.<br>Alto costo computacional de modelos entrenados con imágenes originales, no adecuados para tiempo real.   | Kausar et al., 2023 |
| 12          | LightGBM: A Leading Force in Breast Cancer Diagnosis Through Machine Learning and Image Processing | Dependencia del conjunto de datos BreakHis   | Puede introducir sesgo de selección y dudas sobre la generalización a diferentes poblaciones y condiciones de imagen.   | Kanber et al., 2024 |

**Tabla 1. (Continuación)**

| <b>Ítem</b> | <b>Título del Artículo</b>  | <b>Limitación</b>  | <b>Definición de la Limitación</b>   | <b>Fuente</b>      |
|-------------|---|--|--|--------------------|
| 13          | Potential rapid intraoperative cancer diagnosis using dynamic full-field optical coherence tomography and deep learning: A prospective cohort study in breast cancer patients | Dificultad para diferenciar tipos de tumores<br><br>Generalización a otras poblaciones | Falta de datos de entrenamiento suficientes<br><br>El estudio se centró en un conjunto específico de pacientes, lo que podría limitar la generalización de los resultados. | Zhang et al., 2024 |

Fuente: Elaboración propia

### **III. OBJETIVOS**

#### **3.1. OBJETIVO GENERAL**

- Evaluar los modelos de inteligencia artificial ya existentes para la clasificación de tumores en imágenes de ecografías mamarias, mediante una comparativa de diferentes arquitecturas de redes neuronales convolucionales y la implementación de técnicas de preprocesamiento con el fin de mejorar la precisión en la identificación de tumores en las imágenes de ecografía mamaria.

#### **3.2. OBJETIVOS ESPECÍFICOS:**

- Seleccionar una base de datos diversificada de imágenes de ecografías mamarias, asegurando una clasificación adecuada en normal, benigno y maligno.
- Analizar la influencia del tamaño del lote en la precisión y pérdida del modelo durante su entrenamiento utilizando distintos parámetros.
- Comparar el rendimiento de diferentes arquitecturas de redes neuronales convolucionales, como EfficientNetV2, ResNet50 y VGG16, en la clasificación de tumores mamarios utilizando datos de entrenamiento.
- Validar el modelo de inteligencia artificial adaptado, utilizando un conjunto de datos independiente, analizando su rendimiento mediante la aplicación de métricas como la sensibilidad, especificidad, precisión y F1-score.

## **IV. MÉTODOS**

A continuación, se describirán detalladamente las metodologías y técnicas empleadas en la investigación para lograr la clasificación precisa de tumores mamarios utilizando inteligencia artificial. Se expondrá el proceso de selección y preprocesamiento de las imágenes de ecografías mamarias, las arquitecturas de redes neuronales convolucionales seleccionadas para el estudio, y los procedimientos de entrenamiento y evaluación de los modelos. Además, se presentarán las estrategias de validación y las métricas utilizadas para medir el rendimiento de cada modelo, proporcionando una visión completa y estructurada de la metodología aplicada en este proyecto

### **4.1. ENFOQUE**

La presente investigación adoptó un enfoque experimental en el ámbito de la identificación del cáncer. Utilizó un diseño de aprendizaje supervisado y un estudio transversal que recopiló datos en un único momento. Este enfoque se manifestó en la manipulación controlada de variables para establecer relaciones entre la clasificación precisa de los diferentes tipos de cáncer.

En cuanto al diseño de la investigación de tipo experimental, se basó en el enfoque cuantitativo y en el paradigma deductivo, apoyándose en hipótesis preestablecidas y en la medición de variables. Su aplicación se adhirió al diseño previamente concebido durante su desarrollo, centrándose en la validez, el rigor y el control de la situación de la investigación. Siguiendo a Sampieri, se enfocó en la recolección y análisis de datos cuantitativos. Este enfoque buscaba obtener datos objetivos y verificables que pudieran ser analizados estadísticamente para identificar patrones, relaciones y tendencias en los fenómenos estudiados. En este tipo de diseño, se suelen utilizar instrumentos estandarizados para la recolección de datos y medir con precisión los efectos de las variables independientes sobre las variables dependientes, brindando una comprensión más profunda de los fenómenos investigados (Sampieri et al., 2014).

### **4.2. OBJETIVO DEL MÉTODO**

**Tabla 2.** Objetivo del Método

| <b>Metodología del estudio</b> |   |
|--------------------------------|---|
| Enfoque                        | Cuantitativo  |
| Tipo de estudio                | Transversal   |
| Tipo de diseño                 | Experimental  |
| Técnica instrumentos           | Procesamiento de imágenes con Python/<br>Entrenamiento del modelo Tensorflow /<br>Arquitectura Efficientnetv2 |

Fuente: Elaboración propia

**Tabla 3.** Objetivo del Método

| <b>Método</b>              | <b>Objetivo</b>   | <b>Técnica o Instrumento</b> |
|----------------------------|---|------------------------------|
| Recopilación de datos      | Obtener una base de datos de pacientes con diagnóstico de cáncer de mama diagnosticados por medio de imágenes.  | Open Scholar                 |
| Procesamiento de Imágenes  | Preparar las imágenes para el análisis posterior. Esto incluye la orientación, cambio de tamaño, la conversión a escala de grises, rotación, desenfoque y la adición de ruido. Estos pasos ayudan a resaltar las características importantes y a hacer que el modelo sea más robusto a las variaciones en las imágenes. | Roboflow, Python             |
| Entrenamiento de la red    | Desarrollar y entrenar un modelo de aprendizaje automático que pueda identificar con precisión los diferentes tipos de cáncer. Esto implica la selección de una arquitectura de red adecuada, la definición de una función de pérdida y la optimización de los parámetros de la red utilizando TensorFlow.              | TensorFlow                   |
| Visualización de resultado | Presentar los resultados de manera clara y comprensible para facilitar su interpretación. Esto puede implicar la creación de gráficos y tablas que muestren el rendimiento del modelo, así como la visualización de las predicciones del modelo en las imágenes de ecografía.   | Google Colab, Python         |

Fuente: Elaboración propia

### **4.3. VARIABLES DE INVESTIGACIÓN**

#### 4.3.1. VARIABLE DEPENDIENTES

En el estudio de la clasificación de cáncer de mama mediante inteligencia artificial, la variable dependiente clave es la clasificación del tipo de cáncer. Utilizando un modelo de red neuronal convolucional, se buscó entender cómo los diferentes parámetros y configuraciones del modelo influyen en la capacidad del modelo para clasificar correctamente las imágenes de ecografías mamarias en tres categorías: normal, benigno y maligno.

#### 4.3.2. VARIABLES INDEPENDIENTES

Épocas: Este término se refiere al número de veces que el algoritmo de entrenamiento procesa todo el conjunto de datos de entrenamiento. A mayor número de épocas, el modelo tiene más oportunidades para ajustar sus parámetros y mejorar su capacidad para reconocer patrones en los datos.

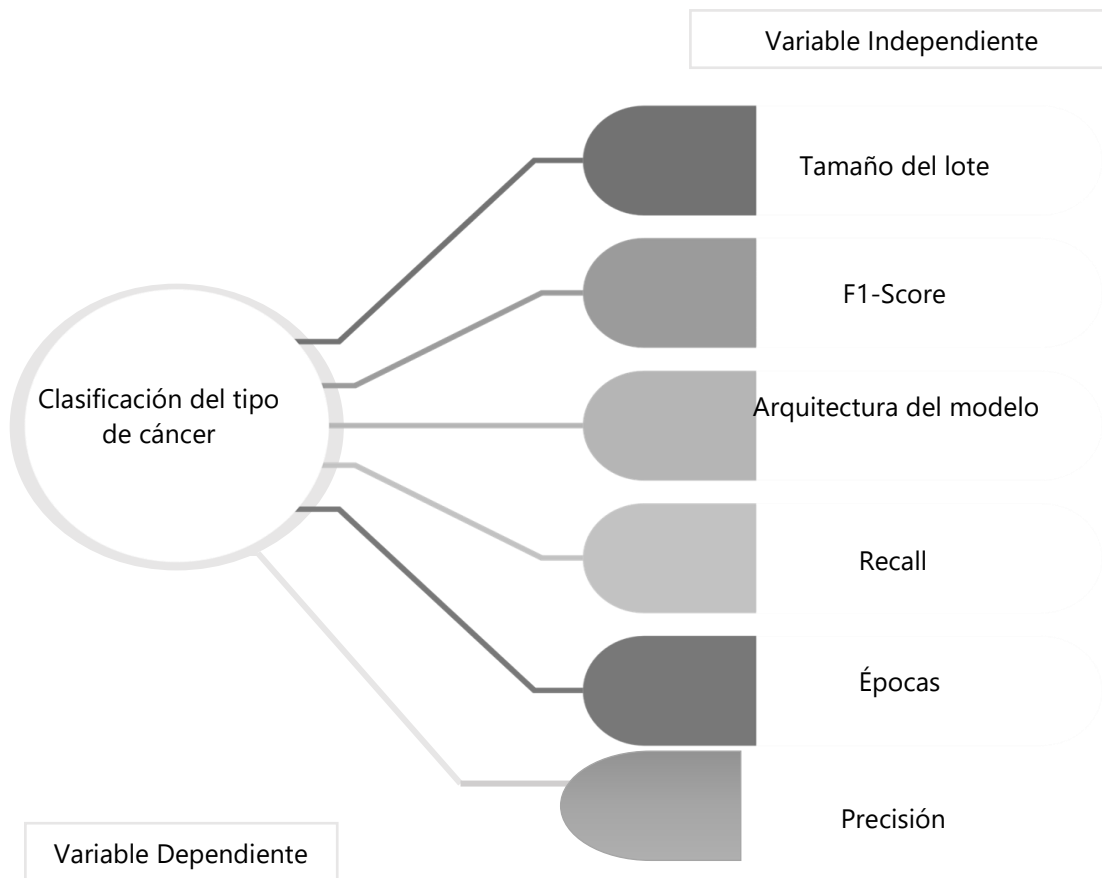
Arquitectura del modelo: Esta variable se refiere a la estructura específica de la red neuronal utilizada, incluyendo el número y tipo de capas (convolucionales, de pooling, densas, etc.), así como la disposición y los hiperparámetros de cada capa. Diferentes arquitecturas pueden afectar significativamente la capacidad del modelo para aprender y generalizar a partir de los datos.

Tamaño del lote: El tamaño del lote se refiere al número de muestras de entrenamiento que se utilizan en una sola iteración del algoritmo de entrenamiento. El tamaño del lote puede influir en la estabilidad y velocidad del entrenamiento, así como en la capacidad del modelo para generalizar.

Precisión: Esta métrica indica la proporción de verdaderos positivos entre todas las predicciones positivas realizadas por el modelo. Una alta precisión significa que el modelo tiene una baja tasa de falsos positivos.

Recall: También conocido como sensibilidad, esta métrica mide la proporción de verdaderos positivos entre todas las muestras que realmente son positivas. Un alto recall indica que el modelo tiene una baja tasa de falsos negativos.

F1 score: Esta métrica es la media armónica de la precisión y el recall, proporcionando una sola puntuación que equilibra ambas métricas. Es especialmente útil cuando hay un desequilibrio entre las clases.



**Ilustración 5.** Variables de Investigación

Fuente: Elaboración propia

## 4.4. RESUMEN DEL PROCEDIMIENTO

### 4.4.1. POBLACIÓN Y MUESTRA

Los registros provinieron de una base de datos que contenía ecografías mamarias de mujeres con edades entre 25 y 75 años, recopilados en 2018. La muestra abarcaba 600 participantes y constaba de 780 imágenes en formato PNG, con dimensiones promedio de 500x500 píxeles. La clasificación de las imágenes estaba dividida en tres categorías: normales, benignas y malignas (Al-Dhabyani et al., 2020). Las imágenes fueron proporcionadas por el Hospital Baheya, ubicado en Egipto, y la base de datos fue gestionada por el Dr. Mohamed Hamed.

### 4.4.2. TÉCNICAS E INSTRUMENTOS APLICADOS

En la presente investigación se emplearon diversas técnicas e instrumentos para alcanzar el objetivo de implementar una red neuronal de inteligencia artificial que pueda detectar tumores en imágenes de ecografías.

Recopilación de Datos: La búsqueda y selección de la base de datos se realizó mediante Google Scholar y Kaggle, utilizando palabras clave específicas para encontrar conjuntos de datos relevantes y de alta calidad.

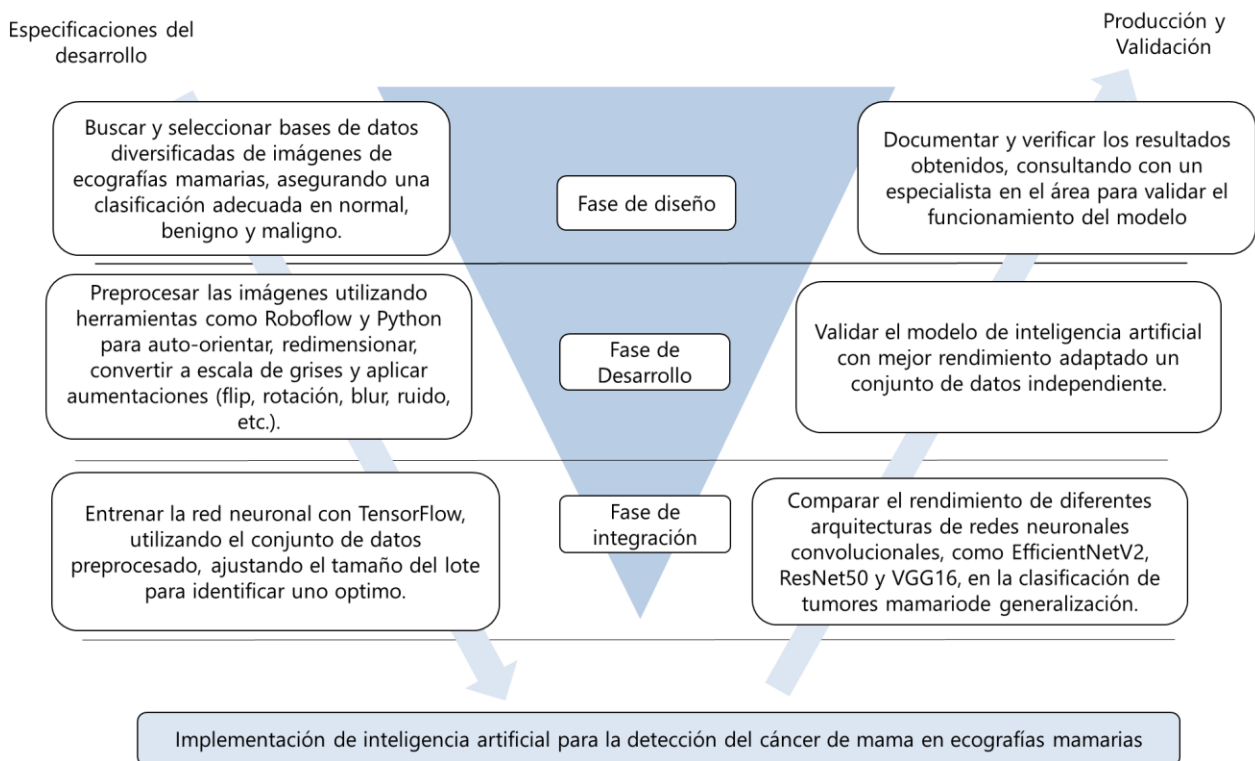
Procesamiento de Imágenes: El procesamiento de las imágenes se llevó a cabo utilizando Roboflow y Python. En esta etapa, se aplicaron técnicas de preprocesamiento como la auto-orientación, el cambio de tamaño de las imágenes a 640x640 píxeles, la conversión a escala de grises y la modificación de clases. Además, se realizaron aumentaciones de datos que incluyeron volteo horizontal y vertical, rotaciones de 90°, rotación aleatoria entre -11° y +11°, cizallamiento horizontal y vertical de  $\pm 15^\circ$ , desenfoque hasta 3.1 píxeles y adición de ruido hasta el 1.92% de los píxeles.

Entrenamiento de la Red: Para el entrenamiento del modelo de red neuronal, se utilizó TensorFlow, una plataforma de aprendizaje automático de código abierto. La arquitectura del modelo fue EfficientNetV2, conocida por su eficiencia y precisión en tareas de clasificación de

imágenes. El entrenamiento se llevó a cabo ajustando parámetros como el número de épocas y el tamaño del lote para optimizar el rendimiento del modelo.

Visualización de Resultados: La visualización de los resultados se realizó utilizando Google Colab y Python, con el apoyo de diversas librerías especializadas como Matplotlib. Estas herramientas permitieron una representación gráfica detallada de las métricas de rendimiento del modelo, como la precisión, el recall y el F1 score, así como la visualización de las predicciones del modelo sobre las imágenes de prueba.

#### 4.4.3. METODOLOGÍA DEL ESTUDIO



**Ilustración 6.** Metodología en V

Fuente: Elaboración propia

Para la presente investigación, se optó por emplear el Modelo en V como metodología de estudio. Esta metodología es adecuada para proyectos que requieren un enfoque estructurado, asegurando que cada fase del desarrollo esté bien definida y validada antes de proceder a la

siguiente etapa. El Modelo en V es una variante del modelo en cascada, que pone un fuerte énfasis en la verificación y validación en cada fase del proyecto. A continuación, se detallan las fases del diseño, desarrollo, integración, validación y el proceso de desarrollo del proyecto.

#### *4.4.3.1. Fase de Diseño*

En la fase de diseño, se realizaron actividades para establecer las bases del proyecto. Primero, se llevó a cabo una búsqueda de bases de datos de imágenes mamográficas disponibles y se evaluaron diversos modelos de aprendizaje automático, seleccionando los más adecuados en función de sus capacidades.

#### *4.4.3.2. Fase de Desarrollo*

La fase de desarrollo involucró el procesamiento de imágenes y el entrenamiento del modelo de inteligencia artificial. Utilizando herramientas como Roboflow y Python, las imágenes fueron preprocesadas, lo que incluyó auto-orientación, redimensionamiento a 640x640 píxeles, conversión a escala de grises, y aplicación de técnicas de aumento de datos como flips horizontales y verticales, rotaciones, shears, blurs y ruido.

#### *4.4.3.3. Fase de Integración*

La fase de integración fue crucial para garantizar la cohesión y funcionalidad del sistema. En esta etapa, se realizaron ajustes del modelo para optimizar su rendimiento en la clasificación de imágenes mamográficas. Posteriormente, el modelo EfficientNetV2 fue entrenado utilizando TensorFlow, ajustando continuamente los parámetros del modelo para asegurar un rendimiento óptimo. Se incluyó pruebas de diferentes configuraciones de tamaño de lote y otros parámetros de entrenamiento, buscando la configuración óptima que maximice la precisión del modelo y minimice la pérdida.

#### *4.4.3.4. Validación de la Integración*

La fase de validación de la integración se centró en comparar el rendimiento de diferentes arquitecturas de redes neuronales convolucionales, como EfficientNetV2, ResNet50 y VGG16, en la detección de tumores mamarios. Utilizando un conjunto de datos de prueba, se aplicaron métricas como la precisión, la pérdida y el tiempo de entrenamiento para identificar el modelo más eficiente y preciso para la clasificación del diagnóstico. Esta comparación ayudó a determinar

cuál de las arquitecturas es más adecuada para la tarea específica de detección de tumores en ecografías mamarias.

#### *4.4.3.5. Validación del Desarrollo*

En la validación del desarrollo, se llevaron a cabo pruebas para asegurar que el modelo se entrenara correctamente y que los resultados fueran consistentes. Se analizaron métricas como precisión, recall y F1 score para evaluar el rendimiento del modelo. Asimismo, se verificó que el modelo pudiera

#### *4.4.3.6. Validación del Diseño*

La validación del diseño se centró en documentar y verificar los resultados obtenidos, consultando con un especialista en el área para validar el funcionamiento del modelo. Se compararon los resultados con otros modelos existentes en la literatura, enfocándose en la precisión diagnóstica de la clasificación en los tipos de cáncer. Esta validación con un especialista ayudó a confirmar la efectividad del modelo y a asegurar que los resultados sean clínicamente relevantes y fiables.

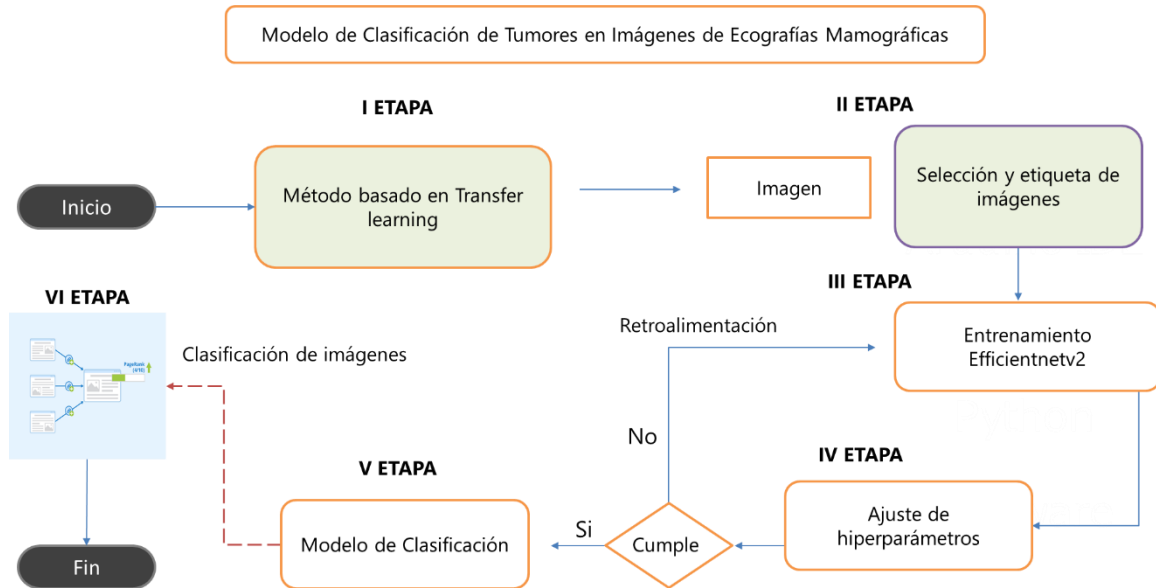
#### *4.4.3.7. Proceso de Desarrollo*

El proceso de desarrollo del prototipo siguió una metodología estructurada. Inicialmente, se configuraron las herramientas y librerías necesarias en Google Colab y Python. Luego, el modelo EfficientNetV2 fue entrenado utilizando TensorFlow, con ajustes continuos para mejorar el rendimiento. Finalmente, se emplearon herramientas de visualización para analizar los resultados y asegurarse de que el modelo cumpliera con los objetivos de precisión y sensibilidad diagnóstica.

### **4.5. IMAGEN DEL PROCESO**

El flujo metodológico empleado en la investigación para detectar cáncer de mama en imágenes de ultrasonido, utilizando técnicas de inteligencia artificial, se representó mediante un proceso integral diseñado para asegurar la precisión y eficacia en la clasificación de imágenes.

Este enfoque se basó en el aprendizaje por transferencia y la arquitectura de EfficientNetV2 (Ilustración 7).



**Ilustración 7.** Resumen del procedimiento

Fuente: Elaboración propia

## 4.6. MÉTRICAS DE EVALUACIÓN Y VALIDACIÓN

La evaluación y validación del modelo de detección de cáncer de mama en imágenes de ultrasonido se basan en varias métricas clave, las cuales permiten determinar la eficacia y precisión del modelo. Estas métricas incluyen la precisión, el recall, el F1 score y la matriz de confusión.

### 4.6.1. PRECISIÓN (ACCURACY)

La precisión es una métrica fundamental que indica el porcentaje de predicciones correctas realizadas por el modelo sobre el total de predicciones. En el contexto de detección de cáncer de mama, la precisión refleja la proporción de imágenes correctamente clasificadas (tanto benignas, malignas como normales) en comparación con el número total de imágenes evaluadas. Una alta precisión sugiere que el modelo está desempeñándose bien en general.

$$Precisión = \frac{TP}{TP + FP}$$

### **Ecuación 1.** Fórmula para el cálculo de precisión

TP = número de verdaderos positivos, FP = número de falsos positivos.

Fuente: (De Carvalho et al., 2023)

#### 4.6.2. RECALL, O SENSIBILIDAD

El recall, o sensibilidad, mide la capacidad del modelo para identificar correctamente las imágenes que contienen tumores. Es la proporción de verdaderos positivos (casos correctamente identificados sobre el total de verdaderos casos positivos). Un alto recall indica que el modelo es eficaz en la detección de casos positivos, minimizando los falsos negativos.

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

### **Ecuación 2.** Fórmula para el cálculo de la sensibilidad

TP = número de verdaderos positivos, FN = número de falsos negativos.

Fuente: (De Carvalho et al., 2023)

#### 4.6.3. F1 SCORE

El F1 score es la media armónica de la precisión y el recall, proporcionando un balance entre ambas métricas. Es especialmente útil cuando se necesita un equilibrio entre precisión y recall, y cuando la distribución de clases es desigual. El F1 score toma en cuenta tanto los falsos positivos como los falsos negativos, ofreciendo una única métrica que refleja tanto la precisión del modelo como su capacidad para detectar correctamente los casos positivos.

$$F1\ Score = \frac{2 \times (Precisión \times Recall)}{(Precisión + Recall)}$$

### **Ecuación 3.** Fórmula para el cálculo del f1-score

Fuente: (De Carvalho et al., 2023)

#### 4.6.4. MATRIZ DE CONFUSIÓN

La matriz de confusión es una herramienta que permite visualizar el rendimiento del modelo en términos de verdaderos positivos, verdaderos negativos, falsos positivos y falsos

negativos. Cada fila de la matriz representa las instancias de una clase real, mientras que cada columna representa las instancias de una clase predicha. La matriz de confusión proporciona una comprensión más detallada de los errores cometidos por el modelo y es útil para identificar si el modelo está confundiendo ciertas clases más que otras.

#### 4.7. CRONOGRAMA DE ACTIVIDADES.

**Tabla 4.** Cronograma de Actividades

| Actividad/Semana  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|----|----|
| Encontrar una base de datos con imágenes de pacientes diagnosticados con cáncer de mama |   |   |   |   |   |   |   |   |   |    |    |
| Selección del modelo de entrenamiento.  |   |   |   |   |   |   |   |   |   |    |    |
| Preprocesamiento de imágenes con herramientas como Roboflow y Python.                   |   |   |   |   |   |   |   |   |   |    |    |
| Entrenamiento inicial de la red neuronal con TensorFlow.                                |   |   |   |   |   |   |   |   |   |    |    |
| Ajuste de parámetros del rendimiento del modelo.  |   |   |   |   |   |   |   |   |   |    |    |
| Validación del rendimiento del modelo.  |   |   |   |   |   |   |   |   |   |    |    |
| Registro de conclusiones sobre la efectividad del modelo.                               |   |   |   |   |   |   |   |   |   |    |    |
| Entrega Avance Final.   |   |   |   |   |   |   |   |   |   |    |    |

Fuente: Elaboración propia

## **V. RESULTADOS**

La presente sección expone los resultados obtenidos a lo largo del desarrollo del proyecto. Expondrá de manera metódica de las actividades realizadas, facilitando la obtención de resultados en cada etapa del proceso. A continuación, se describen los hallazgos de cada fase del proyecto.

### **5.1. BASE DE DATOS**

Se recopilaron imágenes de ecografía mamaria de diversas fuentes, incluyendo Kaggle, Universo Roboflow y Google Scholar, asegurando una diversidad de búsqueda en las muestras. Las imágenes seleccionadas incluyen ecografías mamarias de mujeres de edades comprendidas entre 25 y 75 años.

Para garantizar la calidad de los datos, se establecieron criterios específicos de búsqueda y selección. Primero, se priorizaron estudios que incluyeran diagnósticos confirmados, asegurando que las imágenes representaran correctamente las condiciones normales, benignas y malignas. Además, se seleccionaron únicamente imágenes con una resolución adecuada en formato PNG y libre de artefactos que pudieran interferir en el análisis.

Otro criterio importante fue la accesibilidad y documentación de las bases de datos, ya que aquellas que proporcionaban metadatos detallados y accesibles facilitaron la comprensión y uso de las imágenes en el proyecto. Finalmente, solo se consideraron fuentes reconocidas y confiables, como Kaggle, Roboflow y Google Scholar, para asegurar la validez de los datos recopilados.

### **5.2. MODELO DE APRENDIZAJE**

La estrategia fundamental empleada en este estudio es el Transfer Learning, una técnica que aprovecha el conocimiento adquirido por un modelo en una tarea para mejorar el rendimiento en otra tarea relacionada. Para este propósito, se seleccionó la arquitectura EfficientNetV2.

EfficientNetV2 es una mejora de la arquitectura original EfficientNet, la cual se caracteriza por su capacidad para lograr un alto rendimiento con un número relativamente bajo de

parámetros. Esta arquitectura se basa en un enfoque de escalado compuesto por tres dimensiones: ancho, profundidad y resolución de la red neuronal, lo que permite adaptar el modelo a diferentes recursos computacionales y conjuntos de datos.

Al utilizar EfficientNetV2 como modelo base, se aprovechó el aprendizaje previo realizado en grandes conjuntos de datos de imágenes, como ImageNet, para la tarea específica de clasificación de tumores en ecografías mamarias. Este enfoque permitió acelerar el proceso de entrenamiento y mejorar la generalización del modelo, ya que el modelo pre-entrenado contiene representaciones útiles de características visuales que pueden ser transferidas a la nueva tarea con menor requerimiento de datos de entrenamiento.

### **5.3. PROCESAMIENTO Y PREPARACIÓN DE IMÁGENES**

El conjunto de datos constó de 1877 imágenes de ecografía mamaria. Tras realizar los ajustes necesarios en Roboflow, las imágenes se dividieron en tres subconjuntos: 1644 imágenes para el entrenamiento, representando el 82%; 156 imágenes para la validación, con un 8%; y 77 imágenes para la prueba, con un 4% (Ilustración 8).

En este proyecto, se utilizó Roboflow como una plataforma que facilitó la gestión y el preprocesamiento de datos de imágenes para modelos de aprendizaje automático, con el objetivo de optimizar la clasificación de cáncer de mama en imágenes de ecografía. Roboflow proporcionó herramientas para la anotación, aumentación y preprocesamiento de imágenes, lo cual fue crucial para preparar los datos de manera adecuada antes de alimentar al modelo de aprendizaje automático.

Las imágenes se orientaron automáticamente y se redimensionaron a un tamaño uniforme de 640x640 píxeles para asegurar la consistencia en el entrenamiento del modelo. Esta uniformidad en el tamaño ayudó a evitar problemas relacionados con las dimensiones variables de las imágenes, que pueden afectar el rendimiento del modelo.

Todas las imágenes se convirtieron a escala de grises. Este paso fue importante porque la información relevante en las ecografías mamarias se encuentra principalmente en la intensidad de los píxeles, y convertirlas a escala de grises redujo la complejidad computacional sin perder

información valiosa. La conversión a escala de grises también facilitó la aplicación de técnicas de preprocesamiento y aumentación de datos.

Se aplicaron varias técnicas de aumentación de datos para diversificar aún más el conjunto de datos y mejorar la robustez del modelo. Estas técnicas incluían el volteo horizontal y vertical, rotaciones de 90° en sentido horario y antihorario, así como inversiones completas. Además, se realizaron rotaciones aleatorias entre -11° y +11°, cizalladuras horizontales y verticales de  $\pm 15^\circ$ , desenfoque de hasta 3.1 píxeles y se añadió ruido a hasta 1.92% de los píxeles.

El volteo horizontal y vertical ayudó a incrementar la variabilidad en la orientación de las imágenes, lo que permitió al modelo aprender a reconocer características de los tumores independientemente de su orientación. Las rotaciones de 90° y las inversiones completas introdujeron aún más diversidad, ayudando al modelo a generalizar mejor.

Las rotaciones aleatorias entre -11° y +11° proporcionaron variaciones sutiles en la orientación, lo que hizo al modelo más robusto frente a ligeras inclinaciones en las imágenes de entrada. Las cizalladuras horizontales y verticales de  $\pm 15^\circ$  simulaban deformaciones geométricas que podían ocurrir debido a diferentes posiciones de los pacientes durante las ecografías.

El desenfoque de hasta 3.1 píxeles simuló las variaciones en la calidad de imagen que pueden surgir en la práctica clínica, mientras que la adición de ruido a hasta el 1.92% de los píxeles ayudó al modelo a aprender a manejar y filtrar el ruido, mejorando su capacidad para identificar características relevantes en las imágenes.

**Tabla 5.** Distribución de Datos

| <b>Conjunto de Datos</b> | <b>Cantidad de Imágenes</b> | <b>Porcentaje</b> |
|--------------------------|-----------------------------|-------------------|
| Entrenamiento            | 1644                        | 82%               |
| Prueba                   | 77                          | 4%                |
| Validación               | 156                         | 8%                |

Fuente: Elaboración propia

#### 5.4. ENTRENAMIENTO DE LA RED

El primer paso en el entrenamiento del modelo fue cargar el modelo pre-entrenado EfficientNetV2 utilizando TensorFlow. Esta arquitectura fue seleccionada por su eficiencia en el rendimiento y la cantidad relativamente baja de parámetros que maneja. EfficientNetV2 ya estaba pre-entrenada en grandes conjuntos de datos de imágenes, como ImageNet, lo que permitió aprovechar el conocimiento previo y adaptarlo a la tarea específica de clasificación de tumores en ecografías mamarias.

Se configuraron los parámetros de entrenamiento, tales como el número de épocas (EPOCHS) y el tamaño del lote (batch\_size). Las épocas representan el número de veces que el algoritmo de entrenamiento procesa todo el conjunto de datos, mientras que el tamaño del lote determina la cantidad de datos procesados en cada iteración. Además, se seleccionó una función de pérdida adecuada, en este caso, la pérdida de entropía cruzada categórica, que es comúnmente utilizada en problemas de clasificación. El optimizador elegido fue el algoritmo Adam, conocido por su eficacia en la optimización de redes neuronales, ya que ajusta los pesos de la red durante el entrenamiento para minimizar la función de pérdida.

Para manejar el conjunto de datos de entrenamiento y prueba, se utilizaron generadores de datos de TensorFlow. Estos generadores no solo cargaron los datos, sino que también aplicaron técnicas de aumento de datos, como el volteo horizontal y el zoom, para aumentar la diversidad del conjunto de datos y mejorar la capacidad de generalización del modelo. La aumentación de datos es crucial para simular variaciones en las imágenes que el modelo podría encontrar en un entorno real, ayudando a evitar el sobreajuste.

El modelo se entrenó utilizando el método fit() de TensorFlow. Durante el entrenamiento, los pesos de la red neuronal se optimizaron utilizando el algoritmo Adam, y se minimizó la función de pérdida. El proceso de entrenamiento se realizó en varias épocas, donde cada época consistió en una iteración completa a través de todo el conjunto de datos de entrenamiento. Después de cada época, se evaluó el rendimiento del modelo en el conjunto de datos de validación para monitorizar su progreso y prevenir el sobreajuste. Se calcularon métricas como la precisión y la

pérdida en el conjunto de datos de validación para determinar la eficacia del modelo en la detección de tumores.

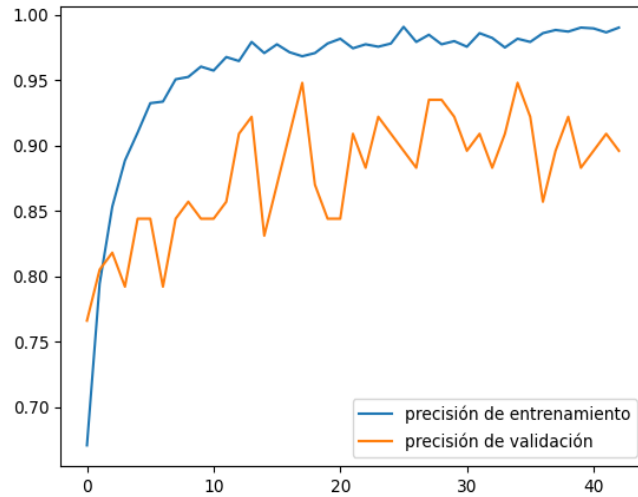
Durante el entrenamiento, se guardó el modelo con la mejor precisión en el conjunto de datos de validación utilizando callbacks de TensorFlow. Estos callbacks ayudaron a guardar el modelo en una ubicación específica para evitar el sobreajuste. Además, se programó una parada temprana en caso de que el modelo no aumentara la precisión de validación después de 25 épocas sin mejora. Esto garantizó que se conservara la mejor versión del modelo para su posterior uso en la clasificación de tumores en imágenes de ecografías mamarias.

#### 5.4.1. INFLUENCIA DEL TAMAÑO DE LOTE EN LA EFICIENCIA DE EFFICIENTNETV2

Para evaluar el impacto del tamaño de lote en la precisión del entrenamiento de EfficientNetV2, se realizaron experimentos con distintos tamaños de lote: 8, 12 y 32. Los resultados obtenidos se resumen a continuación:

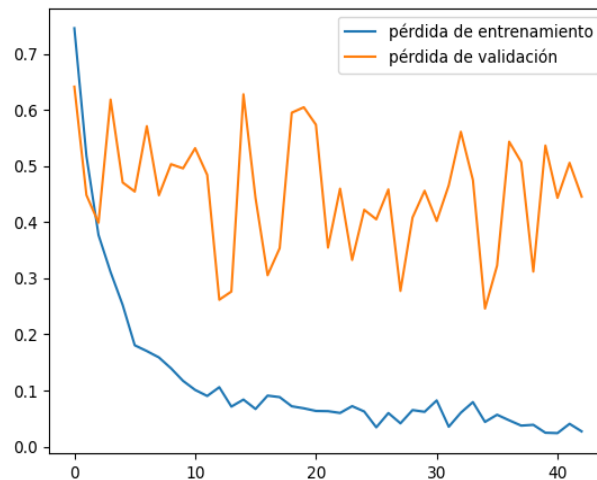
El primer entrenamiento se realizó utilizando un tamaño de lote de 8 y un número máximo de 100 épocas. Sin embargo, el proceso de entrenamiento se detuvo prematuramente en la época 43 debido a la implementación de un criterio de parada temprana, diseñado para detener el entrenamiento una vez que la precisión de validación dejara de mejorar significativamente. La precisión de validación alcanzada fue de 94.80% en la época 18. El modelo mostró una mejora rápida en la precisión durante las primeras épocas y una estabilización relativamente rápida, indicando una convergencia eficiente. A partir de la época 18, comenzaron a observarse fluctuaciones significativas en la precisión, sugiriendo un posible sobreajuste a los datos de entrenamiento (Ilustración 8).

El análisis de las curvas de pérdida reveló que la pérdida de entrenamiento disminuyó de manera constante, alcanzando valores muy bajos hacia las últimas épocas. Sin embargo, la pérdida de validación mostró una tendencia menos consistente, con varias fluctuaciones a lo largo del entrenamiento. Este comportamiento indicó que, aunque el modelo aprendió bien en las primeras épocas, pudo haber comenzado a sobreajustarse a los datos de entrenamiento (Ilustración 9).



**Ilustración 8.** Precisión lote 8 EfficientNetV2

Fuente: Elaboración propia



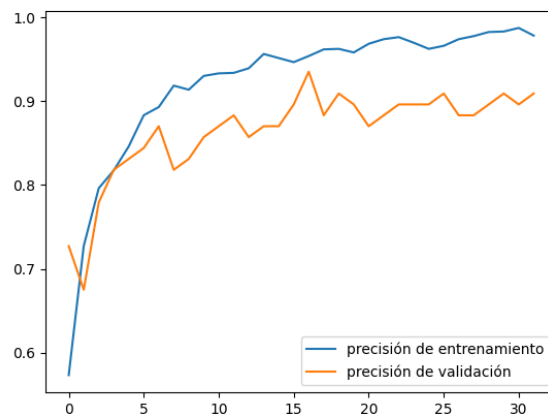
**Ilustración 9.** Perdida lote 8 EfficientNetV2

Fuente: Elaboración propia

Por otro lado, un tamaño de lote de 12 mostró una buena precisión de validación del 93.50%, pero con una mayor variabilidad y menos estabilidad en comparación con el tamaño de lote de 8. Este experimento sugirió que aumentar el tamaño de lote podía introducir más fluctuaciones en el entrenamiento, afectando la consistencia del modelo. Las gráficas correspondientes a este tamaño de lote mostraron más variabilidad y fluctuaciones tanto en la

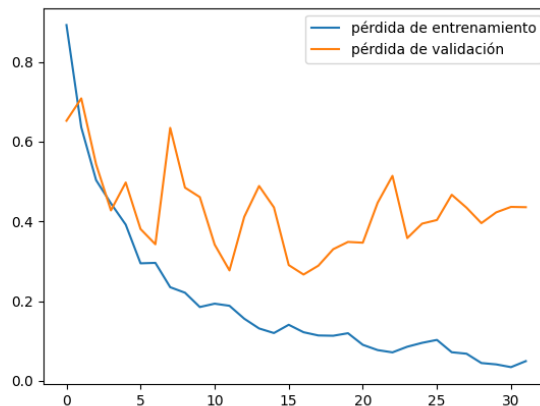
precisión como en la pérdida, indicando un entrenamiento menos estable. Aunque la precisión alcanzada fue aceptable, la mayor variabilidad podría dificultar la replicabilidad y consistencia del modelo en aplicaciones prácticas. El proceso de entrenamiento se detuvo en la época 17 debido al criterio de parada temprana.

La precisión durante el entrenamiento con un tamaño de lote de 12. Se observó una buena precisión, pero con mayor variabilidad y menos estabilidad en comparación con el tamaño de lote de 8. Para la gráfica de pérdida se reflejó fluctuaciones más significativas en la pérdida de validación, indicando un entrenamiento menos estable.



**Ilustración 10.** Precisión lote 12 Efficientnetv2

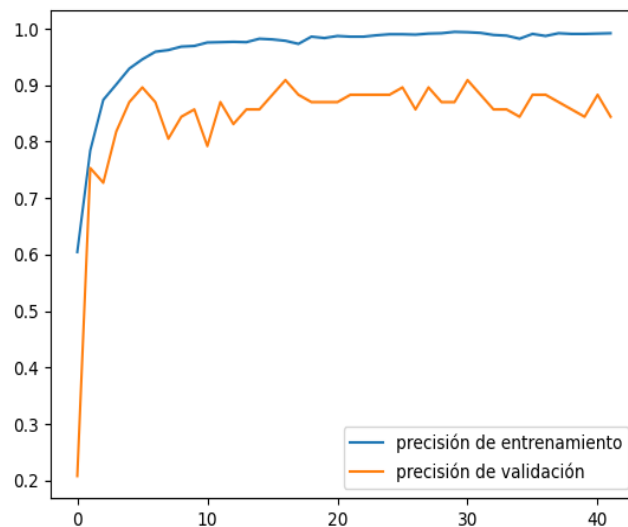
Fuente: Elaboración propia



**Ilustración 11.** Perdida lote 12 Efficientnetv2

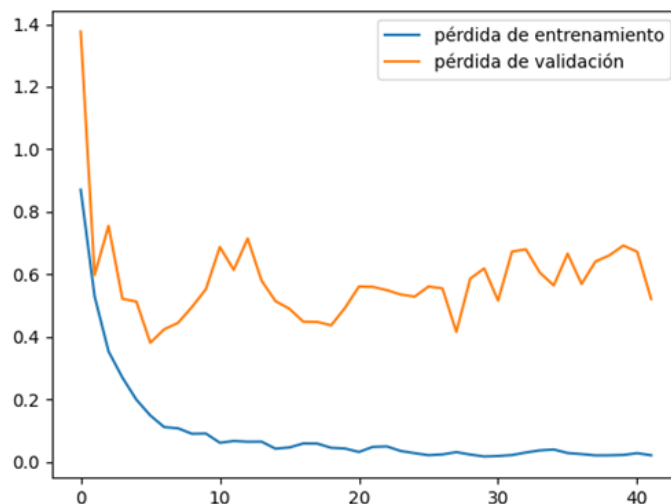
Fuente: Elaboración propia

Finalmente, el entrenamiento con un tamaño de lote de 32 resultó en la menor precisión de validación, alcanzando solo el 90.90%. Además, mostró una mejora mucho más lenta en la precisión, con fluctuaciones significativas. Este comportamiento indicó que un tamaño de lote demasiado grande podía llevar a un aprendizaje ineficiente y a una menor capacidad de generalización del modelo. La mayor fluctuación observada en las métricas de pérdida y precisión sugirió que el modelo tuvo dificultades para estabilizarse y aprender patrones consistentes en los datos de entrenamiento. El entrenamiento se detuvo en la época 42 debido al mismo criterio de parada temprana.



**Ilustración 12.** Precisión lote 32 Efficientnetv2

Fuente: Elaboración propia



**Ilustración 13.** Perdida lote 32 Efficientnetv2

Fuente: Elaboración propia

**Tabla 6.** Resultados de tamaños de lotes

| Tamaño de Lote | Precisión de Validación | Época de Mejor Precisión |
|----------------|-------------------------|--------------------------|
| 8              | 94.80%                  | 18                       |
| 12             | 93.50%                  | 17                       |
| 32             | 90.90%                  | 17                       |

Fuente: Elaboración propia

Estos resultados pusieron de manifiesto que un tamaño de lote más pequeño puede ser el más acertado para evaluar un mejor desempeño, como lo fue con el de tamaño de lote 8, tendió a proporcionar una mayor precisión y estabilidad durante el entrenamiento de EfficientNetV2, en comparación con tamaños de lote más grandes. El tamaño de lote de 8 permitió un ajuste más fino de los parámetros del modelo y una convergencia más rápida y estable. Sin embargo, es importante considerar que un tamaño de lote más pequeño también puede aumentar el tiempo de entrenamiento, ya que realiza más actualizaciones de parámetros por época en comparación con tamaños de lote más grande

#### 5.4.2. COMPARACIÓN EFFICIENTNETV2, RESNET50, VGG16

En esta sección, se compararon tres modelos de redes neuronales convolucionales: EfficientNetV2, ResNet50 y VGG16. Todos los modelos se entrenaron utilizando los mismos parámetros, con un número de épocas (epoch) de 100 y un tamaño de lote (batch size) de 6. Además, se implementó el mecanismo de detención temprana (early stopping) para prevenir el sobreajuste y optimizar el rendimiento de cada modelo.

##### 5.4.2.1. *EfficientNetV2*

Durante el entrenamiento, EfficientNetV2 mostró una eficiencia notable mejorando en cada época y alcanzando su máxima precisión de validación en las primeras épocas. La detención

temprana se activó después de la época 16, cuando ya no se observó una mejora significativa en la precisión de validación.

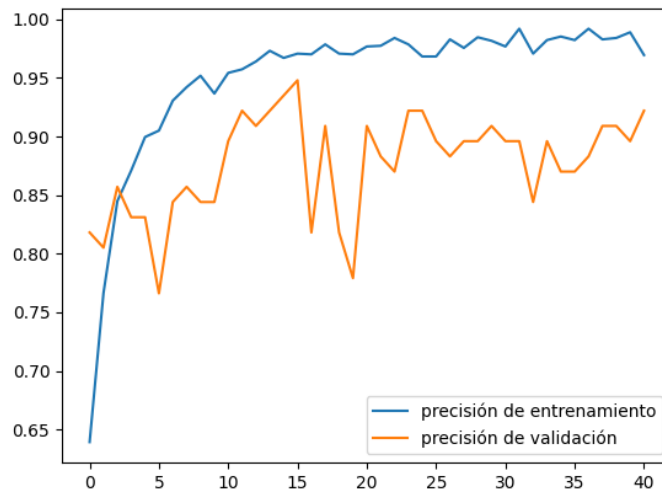
Durante el entrenamiento, EfficientNetV2 mostró una eficiencia notable, mejorando consistentemente en cada época. Este modelo alcanzó su máxima precisión de validación en las primeras épocas, lo que resaltó su capacidad para aprender rápidamente a partir de los datos de entrenamiento. La detención temprana, una técnica utilizada para prevenir el sobreajuste, se activó después de la época 16, cuando ya no se observó una mejora significativa en la precisión de validación. Este mecanismo ayudó a evitar que el modelo continuara entrenándose innecesariamente, lo que podría haber llevado a un sobreajuste y a una disminución en la capacidad de generalización del modelo.

La gráfica de precisión de EfficientNetV2 mostró una rápida estabilización, alcanzando altos niveles de precisión en un tiempo relativamente corto (Ilustración 14). La consistencia de la precisión a lo largo del entrenamiento indicó que el modelo fue capaz de aprender las características importantes de los datos de entrenamiento de manera eficiente, sin grandes fluctuaciones o pérdidas de rendimiento. Esto es crucial para aplicaciones prácticas, donde la estabilidad y la alta precisión del modelo son esenciales para asegurar un rendimiento confiable.

Por otro lado, la gráfica de pérdida de EfficientNetV2 reveló una disminución constante en los valores de pérdida durante el entrenamiento. Una disminución constante en la pérdida es un indicativo de que el modelo estaba mejorando su capacidad para hacer predicciones precisas a medida que se ajustaban los pesos de la red neuronal. La ausencia de fluctuaciones significativas en la pérdida sugirió que el modelo no experimentó problemas de sobreajuste importantes, lo cual es un buen signo de la robustez del modelo.

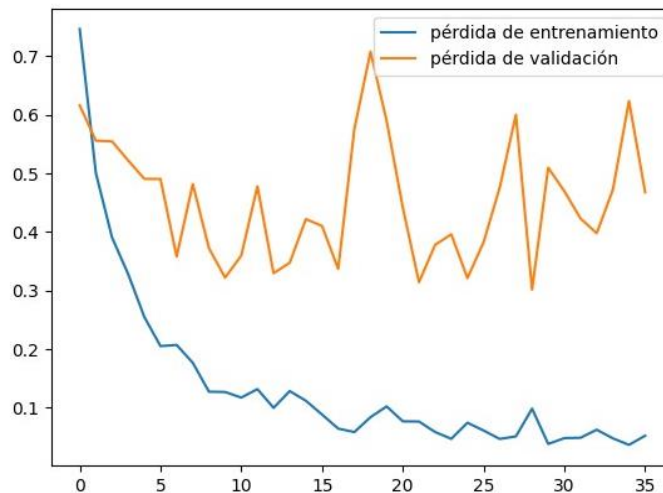
El buen ajuste del modelo, indicado por la disminución constante en la pérdida y la estabilidad en la precisión, reflejó que EfficientNetV2 pudo aprender de los datos de entrenamiento sin memorizar los ejemplos específicos. En cambio, el modelo generalizó bien a nuevas muestras, lo que es una característica esencial para su aplicación en escenarios del mundo real, como la detección de tumores en imágenes de ecografías mamarias.

Además, el tiempo total de entrenamiento de EfficientNetV2 fue de aproximadamente 7 horas, lo que lo hizo más eficiente en términos de tiempo comparado con otros modelos como ResNet50 y VGG16. La capacidad de EfficientNetV2 para alcanzar una alta precisión en un corto período de tiempo y con un número relativamente bajo de parámetros lo posicionó como una opción muy favorable para aplicaciones de diagnóstico médico donde el tiempo y los recursos computacionales son limitados.



**Ilustración 14.** Precisión Efficientnetv2

Fuente: Elaboración propia



**Ilustración 15.** Perdida Efficientnetv2

Fuente: Elaboración propia

#### 5.4.2.2. *ResNet50v2*

El modelo ResNet50v2 también se entrenó con 100 épocas y un tamaño de lote de 6. Sin embargo, el rendimiento de ResNet50 fue un poco inferior en comparación con EfficientNetV2. A lo largo del entrenamiento, la precisión de validación de ResNet50 se estabilizó y el mecanismo de detención temprana se activó en la época 60. Este mecanismo de detención temprana se implementó para prevenir el sobreajuste, permitiendo que el entrenamiento se detuviera cuando ya no se observaban mejoras significativas en la precisión de validación.

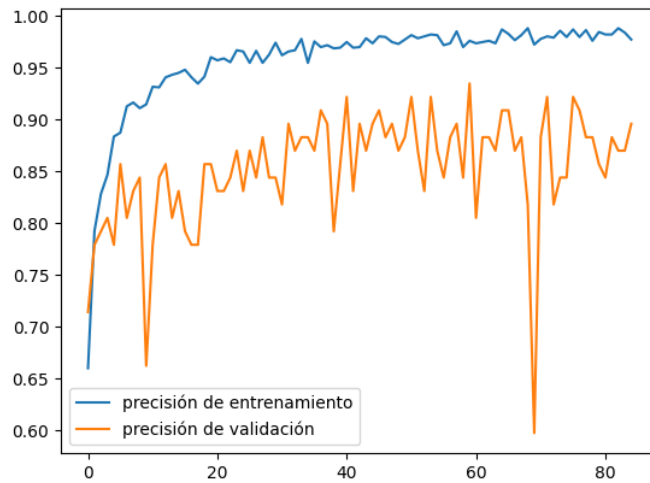
La gráfica de precisión de ResNet50 mostró fluctuaciones notables y una mejora más lenta en comparación con EfficientNetV2 (Ilustración 16). Estas fluctuaciones sugirieron que el modelo tuvo más dificultades para aprender de los datos de entrenamiento y pudo estar más propenso al sobreajuste. A diferencia de EfficientNetV2, que mostró una estabilización rápida y consistente en la precisión, ResNet50 presentó variaciones considerables, lo que indicó una menor estabilidad en su proceso de aprendizaje.

La gráfica de pérdida de ResNet50 reflejó estas dificultades, presentando fluctuaciones significativas a lo largo del entrenamiento (Ilustración 17). Estas fluctuaciones en la pérdida indicaron que el modelo experimentó problemas de sobreajuste, donde el modelo comenzó a ajustar demasiado a los datos de entrenamiento en lugar de aprender patrones generales que pudieran aplicarse a nuevos datos. Esto se tradujo en una menor capacidad de generalización en comparación con EfficientNetV2.

Aunque ResNet50 alcanzó su mejor precisión alrededor de las 24 horas de entrenamiento, su tiempo total de entrenamiento fue de 32 horas, considerablemente mayor que el de EfficientNetV2. Este tiempo de entrenamiento más largo sugiere que ResNet50 necesitó más tiempo para intentar alcanzar un rendimiento óptimo, pero aun así no logró superar la eficiencia de EfficientNetV2.

La mayor duración del entrenamiento y las fluctuaciones en las métricas de rendimiento indicaron que ResNet50, a pesar de ser un modelo robusto, no fue tan efectivo como EfficientNetV2 en términos de velocidad de convergencia y estabilidad del aprendizaje. Estas

características son cruciales en aplicaciones prácticas, especialmente en el ámbito médico, donde la precisión, la estabilidad y el tiempo de entrenamiento son factores determinantes.



**Ilustración 16.** Precisión ResNet50v2

Fuente: Elaboración propia



**Ilustración 17.** Perdida ResNet50v2

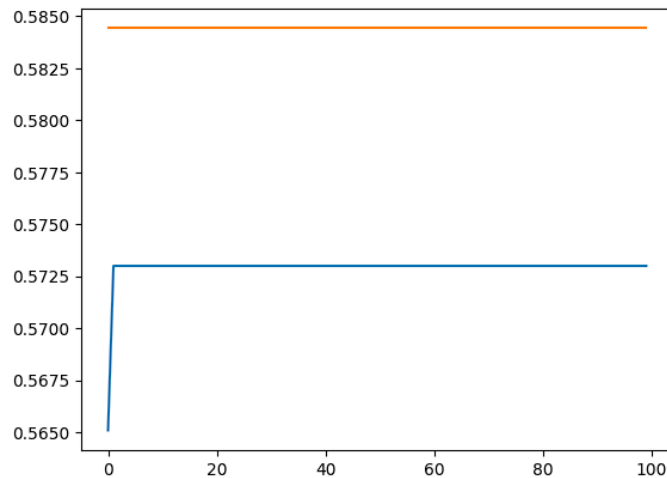
Fuente: Elaboración propia

### 5.4.2.3. VGG16

El modelo alcanzó su máxima precisión de validación desde el principio, no mostro cambios durante el entrenamiento momento en el cual la detención temprana no se activó debido a la falta de mejoras adicionales en la precisión de validación. El modelo VGG16 alcanzó su máxima

precisión de validación desde el principio y no mostró cambios significativos durante el entrenamiento. La detención temprana no se activó debido a la falta de mejoras adicionales en la precisión de validación. Esto sugirió que VGG16 no aprovechó adecuadamente los datos de entrenamiento para mejorar su rendimiento.

La precisión de validación de VGG16 se estabilizó tempranamente sin mejoras significativas posteriores, alcanzando una precisión de validación del 58.44%. Esta estabilización temprana indicó que el modelo había llegado rápidamente a su capacidad máxima de aprendizaje a partir de los datos de entrenamiento proporcionados. Sin embargo, esta capacidad máxima fue considerablemente más baja en comparación con los otros modelos, lo que reflejó una limitación en la habilidad de VGG16 para aprender y generalizar a partir de los datos (Ilustración 18).

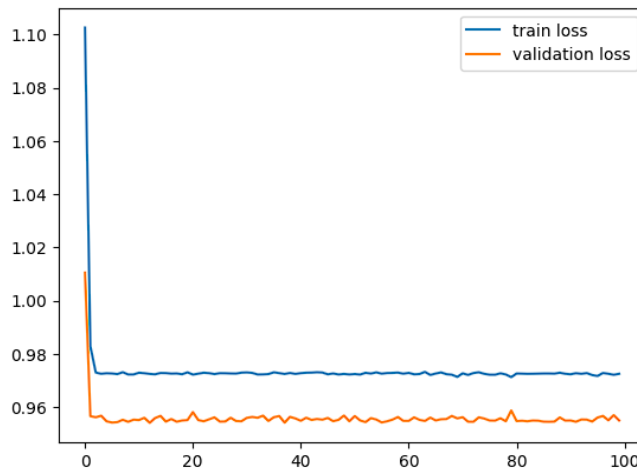


**Ilustración 18.** Precisión VGG16

Fuente: Elaboración propia

La gráfica de pérdida de VGG16 también mostró una estabilización temprana, indicando un rendimiento no óptimo y una menor capacidad de generalización en comparación con EfficientNetV2 y ResNet50. La pérdida se estabilizó sin las típicas disminuciones continuas que se observan en modelos que siguen aprendiendo y mejorando a lo largo del entrenamiento. Este patrón de estabilización temprana sugirió que VGG16 dejó de mejorar casi inmediatamente después del inicio del entrenamiento, lo cual es una señal de que el modelo no pudo extraer más información útil de los datos de entrenamiento (Ilustración 19).

En comparación con EfficientNetV2 y ResNet50, VGG16 demostró ser menos efectivo para la tarea de detección de tumores en imágenes de ecografías mamarias. La falta de mejoras significativas en la precisión y la estabilización temprana de la pérdida indicaron que VGG16 no fue capaz de adaptarse y aprender de manera efectiva a partir de los datos de entrenamiento. Esto hizo que VGG16 fuera menos adecuado para aplicaciones prácticas que requieren alta precisión y buena capacidad de generalización, como el diagnóstico médico automatizado.



**Ilustración 19.** Perdida VGG16

Fuente: Elaboración propia

**Tabla 7.** Resultados de modelos

| Métrica                        | EfficientNetV2                         | ResNet50                             | VGG16                               |
|--------------------------------|--|--------------------------------------|-------------------------------------|
| Época de Mejor Precisión       | 16                                     | 60                                   | N/A                                 |
| Comportamiento de la Precisión | Rápida estabilización y alta precisión | Fluctuaciones notables, lenta mejora | Estabilización temprana sin mejoras |
| Comportamiento de la Pérdida   | Disminución constante y baja pérdida   | Fluctuaciones notables               | Estabilización temprana sin mejoras |
| Precisión de Validación        | 94.81%                                 | 93.51%                               | 58.44%                              |

Fuente: Elaboración propia

La implementación de los mismos parámetros de entrenamiento y el uso del mecanismo de detención temprana permitieron una evaluación justa de cada modelo. EfficientNetV2 mostró una rápida mejora y estabilización en la precisión, alcanzando una precisión de validación del 94.81%. Esto indicó que el modelo aprendió de manera eficiente y alcanzó un alto rendimiento. La baja fluctuación en la precisión sugirió que el modelo obtuvo un buen aprendizaje. La gráfica de pérdida de EfficientNetV2 mostró una disminución constante, lo que indicó un buen ajuste del modelo sin signos significativos de sobreajuste. Además, el tiempo total de entrenamiento fue de alrededor de 7 horas, lo que hizo que EfficientNetV2 fuera más eficiente en términos de tiempo en comparación con los otros modelos.

ResNet50 mostró una mejora más lenta y con fluctuaciones notables en la precisión, alcanzando una precisión de validación del 93.51%. Esto sugirió que el modelo tuvo más dificultades para aprender y pudo estar más propenso al sobreajuste. La gráfica de pérdida de ResNet50 presentó fluctuaciones significativas, lo que pudo indicar problemas de sobreajuste y una menor capacidad de generalización en comparación con EfficientNetV2. Aunque ResNet50 alcanzó su mejor precisión alrededor de las 24 horas de entrenamiento, su tiempo total de entrenamiento fue de 32 horas, considerablemente mayor.

VGG16 mostró una estabilización temprana en la precisión sin mejoras significativas posteriores, alcanzando una precisión de validación del 58.44%. Esto sugirió que el modelo no aprovechó adecuadamente los datos de entrenamiento para mejorar su rendimiento. La gráfica de pérdida de VGG16 también mostró una estabilización temprana, indicando un rendimiento no óptimo y una menor capacidad de generalización en comparación con los otros modelos.

#### 5.4.3. ANÁLISIS DEL TIEMPO DE ENTRENAMIENTO

El tiempo de entrenamiento es un factor crucial para la aplicabilidad práctica de los modelos en entornos clínicos. Se realizó una comparación detallada del tiempo de entrenamiento para los modelos EfficientNetV2, ResNet50v2 y VGG16.

**Tabla 8.** Resultados de modelos

| <b>Modelo</b>  | <b>Tiempo Total de Entrenamiento</b> |
|----------------|--------------------------------------|
| EfficientNetV2 | 7 horas                              |
| ResNet50v2     | 32 horas                             |
| VGG16          | 35 minutos                           |

Fuente: Elaboración propia

EfficientNetV2 no solo mostró una mejor precisión y estabilidad en el rendimiento, sino que también requirió significativamente menos tiempo de entrenamiento en comparación con ResNet50v2. Esta eficiencia en tiempo hace que EfficientNetV2 sea una opción más práctica para implementaciones en la vida real donde el tiempo y los recursos computacionales son limitados.

ResNet50v2, aunque efectivo en términos de precisión, necesitó considerablemente más tiempo para alcanzar su mejor rendimiento. Con un tiempo total de entrenamiento de 32 horas y un tiempo promedio por época de 20 minutos, ResNet50v2 fue menos eficiente comparado con EfficientNetV2. Las fluctuaciones observadas en su precisión y pérdida indicaron que el modelo tuvo más dificultades para estabilizar su aprendizaje, lo que resultó en una mayor duración del entrenamiento.

VGG16, por otro lado, completó su proceso de entrenamiento en alrededor de 35 minutos. Sin embargo, esta rapidez no se tradujo en un rendimiento efectivo. La falta de mejoras significativas en la precisión y la estabilización temprana en la pérdida sugirieron que VGG16 no aprovechó adecuadamente el tiempo de entrenamiento para lograr un ajuste óptimo y una mayor precisión. A pesar de su rapidez, la menor precisión y capacidad de generalización limitaron su utilidad en comparación con EfficientNetV2 y ResNet50v2.

## **5.5. VALIDACIÓN DEL MODELO CON CONJUNTO DE PRUEBA INDEPENDIENTE**

Para evaluar la capacidad de generalización y precisión del modelo en la práctica clínica, se realizó una validación independiente utilizando un conjunto de datos de prueba. Este conjunto

de datos se mantuvo completamente separado durante el entrenamiento para asegurar una evaluación imparcial del rendimiento del modelo.

El modelo que se utilizó para esta validación fue EfficientNetV2, que mostró la mejor precisión de validación durante el entrenamiento, específicamente el guardado en la época 16 con un tamaño de lote de 6. Este modelo se cargó y probó con las imágenes del conjunto de datos de prueba, contenidas en una carpeta independiente de test.

El conjunto de datos de prueba incluía un total de 77 imágenes. Estas imágenes se dividieron en tres categorías: 45 imágenes benignas, 20 imágenes malignas y 12 imágenes normales. Esta distribución permitió evaluar el rendimiento del modelo en distintas clases, proporcionando una visión completa de su capacidad de detección.

Los resultados obtenidos se evaluaron utilizando métricas estándar de clasificación para determinar la precisión, sensibilidad (recall), F1-score y la matriz de confusión. Estas métricas permitieron una comprensión detallada del desempeño del modelo en un contexto clínico, evaluando su capacidad para detectar correctamente los tumores en las imágenes de ecografías mamarias.

## **5.6. DOCUMENTACIÓN DE LOS RESULTADOS OBTENIDOS**

La validación del modelo utilizando dos arquitecturas de redes neuronales convolucionales: EfficientNetV2, ResNet50. Estos modelos se entrenaron y evaluaron utilizando un conjunto de datos independiente para determinar su capacidad de generalización y precisión de imágenes no conocidas por el modelo, específicamente en la detección de cáncer de mama en imágenes de ecografía.

El análisis de los resultados muestra que EfficientNetV2, con 16 épocas de entrenamiento, logró una precisión de validación ligeramente superior (0.95) en comparación con ResNet50V2 (0.94) entrenado por 60 épocas. Las métricas de recall y f1-score también indican un mejor rendimiento general de EfficientNetV2. Se evaluó el modelo con el conjunto de datos de la carpeta "test" la cual contenía: 45 imágenes de tipo benigno, 20 imágenes tipo maligno, y normal que

contenía 12 imágenes. En cada clase se realizaron las evaluaciones para obtener las métricas de cómo se desarrolla el modelo.

**Tabla 9.** Métricas de Clasificación EfficientNetV2 (16 épocas, tamaño de lote 6)

| Clase     | Precisión | Recall | F1-Score | Soporte |
|-----------|-----------|--------|----------|---------|
| Benigno   | 0.94      | 0.98   | 0.96     | 45      |
| Maligno   | 0.94      | 0.85   | 0.89     | 20      |
| Normal    | 1.00      | 1.00   | 1.00     | 12      |
| Precisión | 0.95      | -      | -        | 77      |

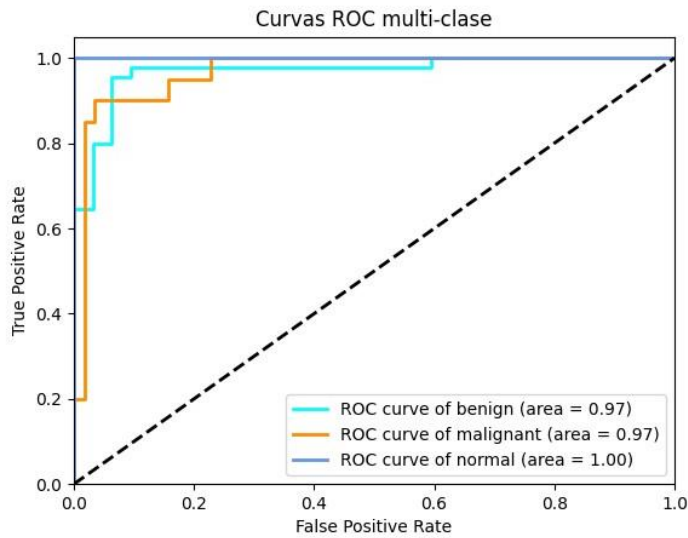
Fuente: Elaboración propia

Para evaluar y comparar la capacidad de discriminación de los modelos EfficientNetV2 y ResNet50v2 en la detección de cáncer de mama, se generaron las curvas ROC (Receiver Operating Characteristic) multi-clase para ambos modelos. Las curvas ROC permiten visualizar el rendimiento de los modelos en términos de su tasa de verdaderos positivos frente a la tasa de falsos positivos para cada categoría.

En la gráfica de curvas ROC multi-clase para EfficientNetV2 (Ilustración 24), se observa el rendimiento del modelo en la clasificación de las imágenes en las clases benigna, maligna y normal. La curva ROC de la clase benigna tiene un área bajo la curva (AUC) de 0.97, lo que indica un excelente rendimiento en la clasificación de imágenes benignas. La curva ROC de la clase maligna también muestra un AUC de 0.97, reflejando una alta capacidad del modelo para detectar imágenes malignas con precisión. Por último, la curva ROC de la clase normal tiene un AUC de 1.00, demostrando que el modelo clasifica las imágenes normales sin ningún error.

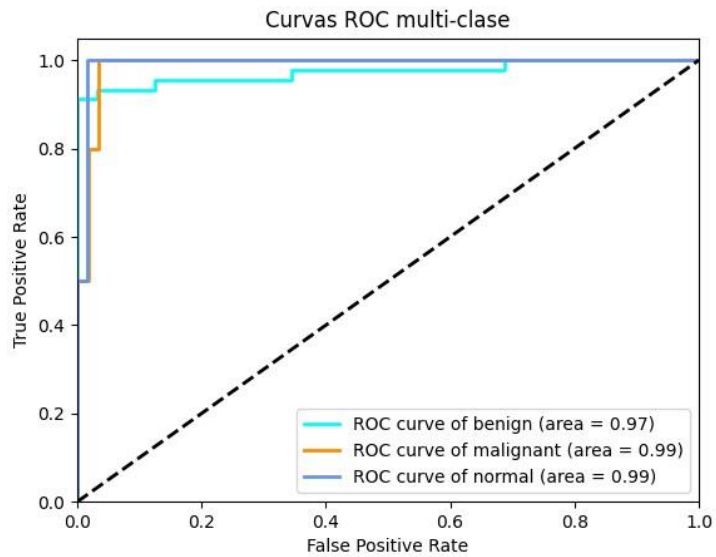
Por otro lado, en la gráfica de curvas ROC multi-clase para ResNet50v2 (Ilustración 25), se observa que la curva ROC de la clase benigna tiene un AUC de 0.97, indicando un rendimiento muy bueno en la clasificación de imágenes benignas. La curva ROC de la clase maligna muestra un AUC de 0.99, reflejando una alta precisión del modelo en la detección de imágenes malignas.

La curva ROC de la clase normal también presenta un AUC de 0.99, demostrando una alta capacidad de clasificación para imágenes normales.



**Ilustración 20.** Curva ROC EfficientNetV2

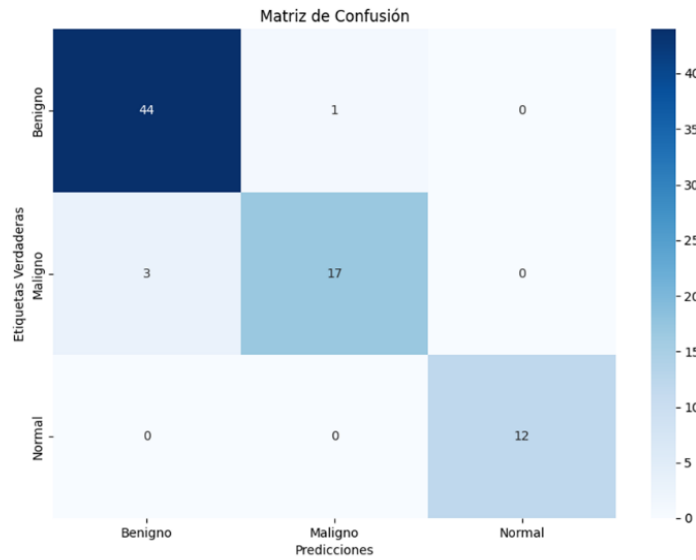
Fuente: Elaboración propia



**Ilustración 21.** Curva ROC ResNet50v2

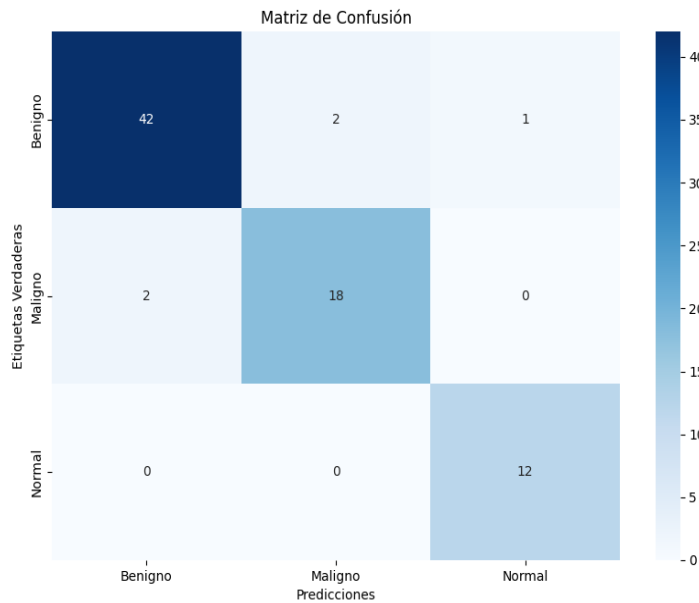
Fuente: Elaboración propia

Las matrices de confusión muestran que EfficientNetV2 tuvo un menor número de predicciones incorrectas en la clase benigna en comparación con ResNet50V2, lo que refuerza la observación de una mayor precisión en el modelo de EfficientNetV2 (Ilustración 21).



**Ilustración 22.** Matriz de Confusión EfficientNetv2

Fuente: Elaboración propia



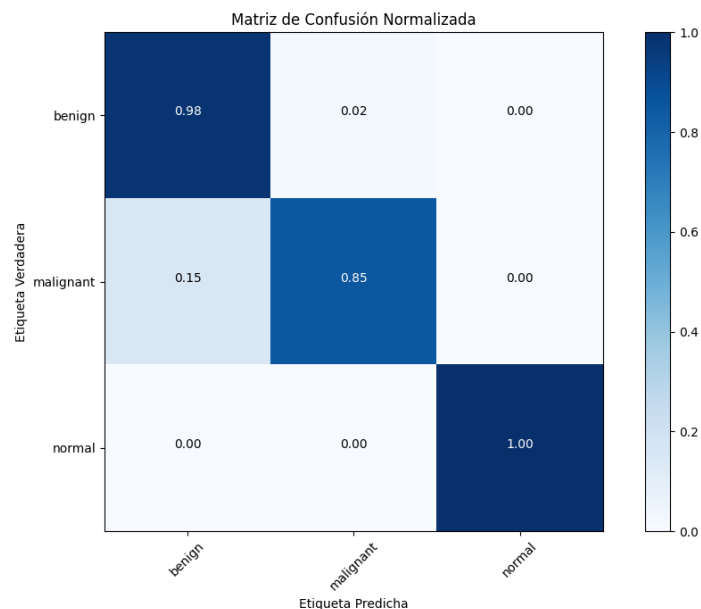
**Ilustración 23.** Matriz de Confusión ResNetv2

Fuente: Elaboración propia

La matriz de confusión normalizada es una versión de la matriz de confusión en la que cada valor se divide por el total de instancias en la clase verdadera correspondiente. Esto convierte los valores absolutos en proporciones o porcentajes, lo que permite una comparación más clara del rendimiento del modelo en diferentes clases, especialmente cuando las clases están desbalanceadas.

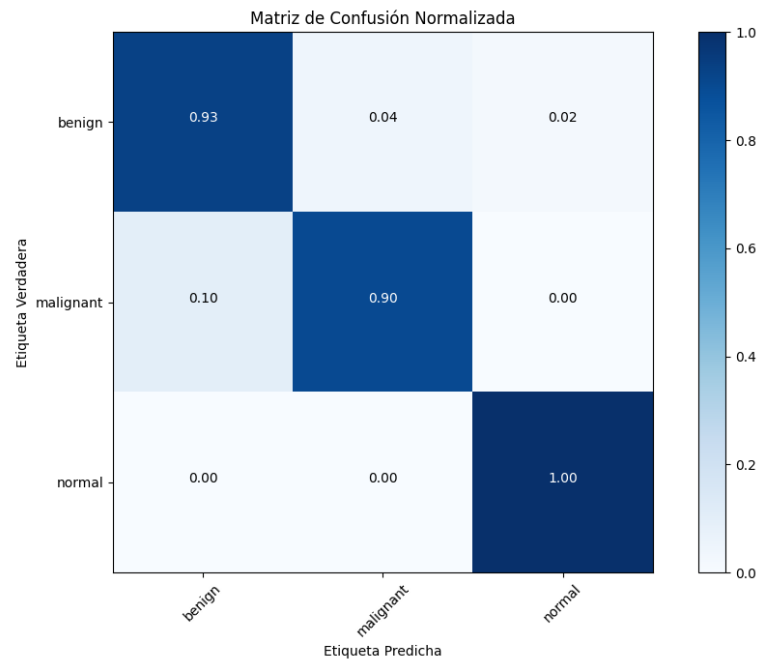
Para proporcionar una visión más clara del rendimiento del modelo, se generó una matriz de confusión normalizada que muestra la precisión de las predicciones de EfficientNetV2 en cada categoría. En la matriz de confusión normalizada, se observa que el modelo logró una precisión del 98% para la clase benigna, 85% para la clase maligna y 100% para la clase normal, también se generó una matriz de confusión normalizada para las predicciones de ResNet50v2 en cada categoría.

El modelo logró una precisión del 93% para la clase benigna, 90% para la clase maligna y 100% para la clase normal. Estos resultados indican que ResNet50v2 tiene una buena capacidad de generalización y precisión en la detección de tumores en imágenes de ecografías mamarias, aunque presenta una ligera desventaja en comparación con EfficientNetV2, especialmente en la clase benigna.



**Ilustración 24.** Matriz de Confusión Normalizada EfficientNetV2

Fuente: Elaboración propia

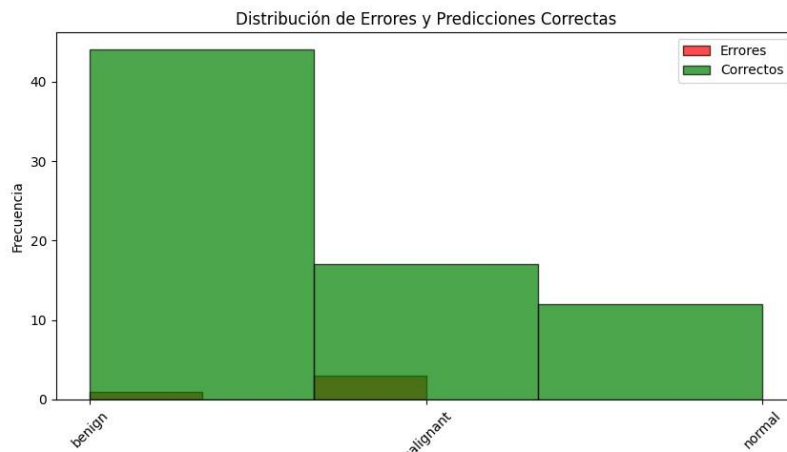


**Ilustración 25.** Matriz de Confusión Normalizada ResNet50v2

Fuente: Elaboración propia

La matriz de confusión logra mostrar el rendimiento de los errores y predicciones para lograr una mejor representación del rendimiento del modelo EfficientNetV2, se generó una gráfica que muestra la distribución de errores y predicciones correctas para cada categoría. En la gráfica de distribución de errores y predicciones correctas (Ilustración 23), se observa que el modelo logró una alta cantidad de predicciones correctas en las clases benigna y normal, con 44 y 12 predicciones correctas, respectivamente. Sin embargo, en la clase maligna, el modelo tuvo una mayor proporción de errores, con 17 predicciones correctas y algunos errores significativos. Estos resultados resaltan la necesidad de mejorar la precisión del modelo en la clase maligna, aunque muestra una fuerte capacidad de generalización en las clases benigna y normal.

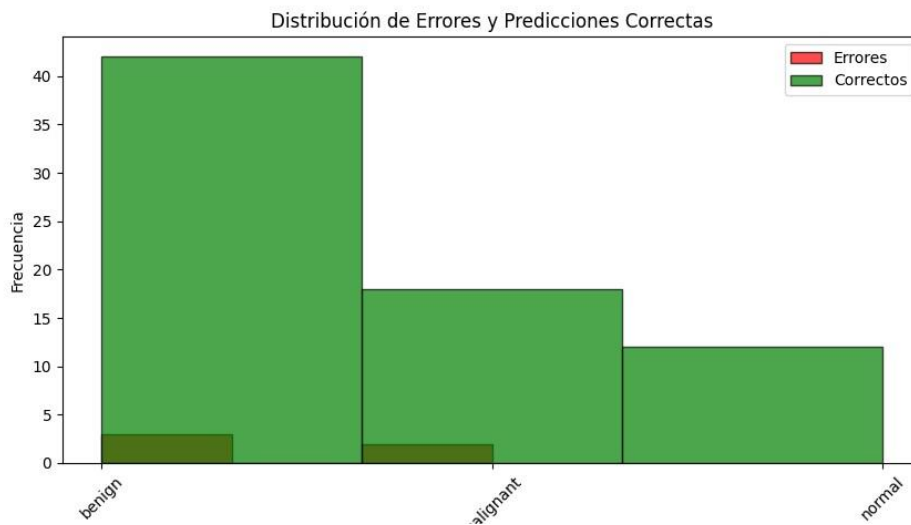
La gráfica proporciona una representación visual clara de la eficiencia del modelo y las áreas donde aún se requieren mejoras para optimizar su rendimiento en la detección de tumores en imágenes de ecografías mamarias.



**Ilustración 26.** Distribución de errores y predicciones EfficientNetv2

Fuente: Elaboración propia

En la gráfica de distribución de errores y predicciones correctas para el modelo de Resnet50v2, se observa que el modelo logró una alta cantidad de predicciones correctas en las clases benigna y normal, con 42 y 12 predicciones correctas, respectivamente. Sin embargo, en la clase maligna, el modelo tuvo una proporción mayor de errores, con 18 predicciones correctas y algunos errores significativos.



**Ilustración 27.** Distribución de errores y predicciones ResNet50v2

Fuente: Elaboración propia

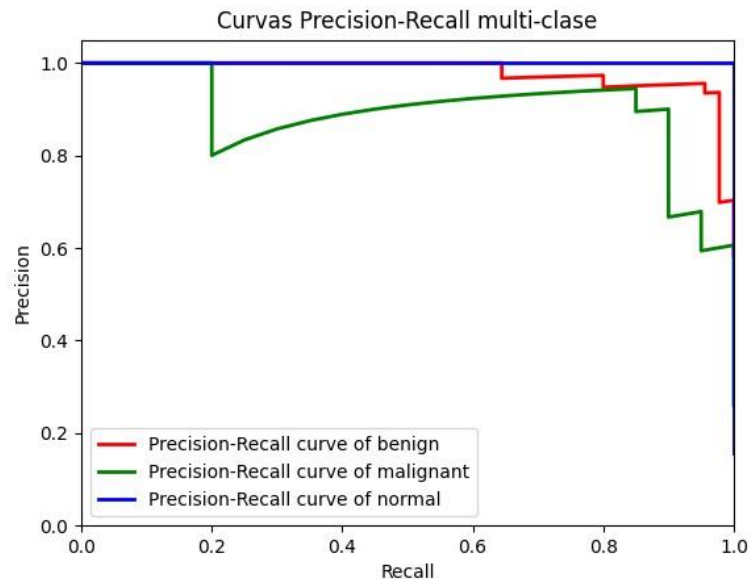
Las métricas de evaluación del modelo ResNet50v2 entrenado con 16 épocas y un tamaño de lote de 6, fueron las mismas del modelo anterior incluyeron la precisión, recall y f1-score para cada clase (benigno, maligno y normal), así como las métricas promedio que proporcionan una visión general del rendimiento del modelo en el conjunto de datos de validación.

**Tabla 10.** Métricas de Clasificación ResNet50v2 (16 épocas, tamaño de lote 6)

| Clase     | Precisión | Recall | F1-Score | Soporte |
|-----------|-----------|--------|----------|---------|
| Benigno   | 0.95      | 0.93   | 0.94     | 45      |
| Maligno   | 0.90      | 0.90   | 0.90     | 20      |
| Normal    | 0.92      | 1.00   | 0.96     | 12      |
| Precisión | 0.94      | -      | -        | 77      |

Fuente: Elaboración propia

La comparación de las curvas Precision-Recall multi-clase entre EfficientNetV2 y ResNet50v2 proporciona una visión clara del rendimiento de ambos modelos en términos de precisión y recall para cada clase (benigno, maligno y normal).

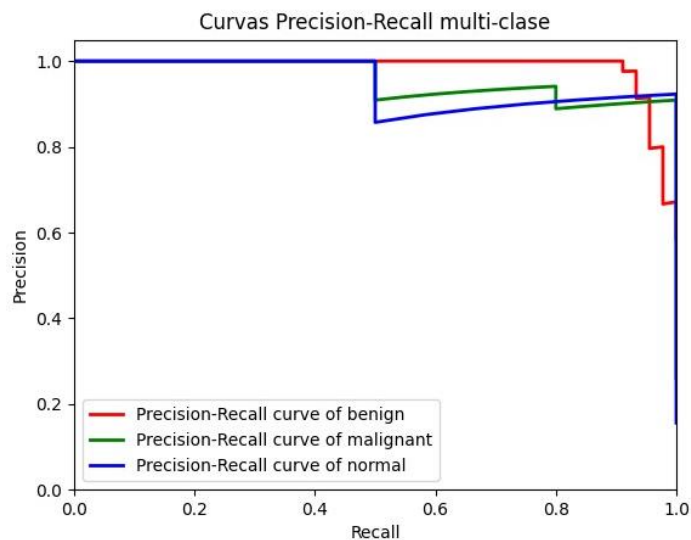


**Ilustración 28.** Curva Precision-Recall EfficientNetV2

Fuente: Elaboración propia

En la gráfica de curvas Precision-Recall multi-clase para EfficientNetV2, se observa que, para la clase benigna, la curva muestra una alta precisión que disminuye ligeramente a medida que aumenta el recall. Esto indica que el modelo puede identificar correctamente la mayoría de las imágenes benignas, aunque con algunos errores a medida que trata de capturar todos los casos positivos. Para la clase maligna, la curva sugiere que el modelo mantiene una precisión razonable mientras incrementa el recall, pero puede perder precisión a medida que trata de identificar todos los casos malignos. En la clase normal, la curva se mantiene alta tanto en precisión como en recall, indicando un excelente rendimiento en la identificación de imágenes normales.

Por otro lado, en la gráfica de curvas Precision-Recall para ResNet50v2, la clase benigna muestra una alta precisión, pero con una ligera caída cuando el recall aumenta. Esto indica que ResNet50v2 puede identificar correctamente la mayoría de las imágenes benignas, aunque puede haber más errores en comparación con EfficientNetV2 cuando se trata de maximizar el recall. La curva de precisión-recall para la clase maligna muestra un comportamiento similar al de EfficientNetV2, con buena precisión y recall, aunque con una ligera disminución en la precisión a medida que se incrementa el recall. Para la clase normal, la curva muestra una alta precisión y recall, similar a EfficientNetV2, indicando un buen rendimiento en la identificación de imágenes normales.



**Ilustración 29.** Curva Precision-Recall ResNet50v2

Fuente: Elaboración propia

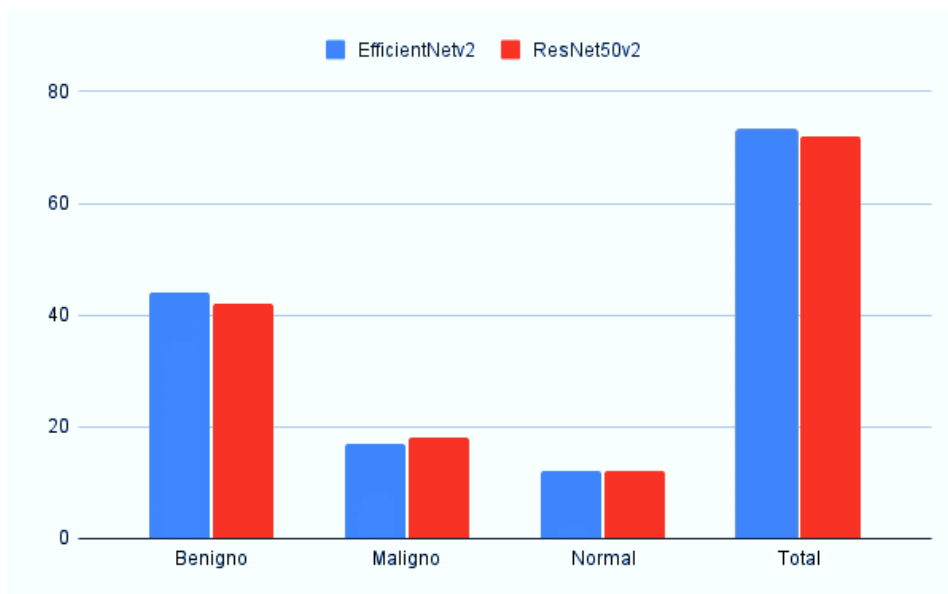
En comparación, las curvas para la clase benigna en ambos modelos muestran alta precisión con una ligera disminución al aumentar el recall. EfficientNetV2 parece tener una caída menos pronunciada en comparación con ResNet50v2, sugiriendo un mejor manejo de las imágenes benignas. Las curvas para la clase maligna muestran que ambos modelos mantienen buena precisión con una ligera disminución al aumentar el recall. Sin embargo, EfficientNetV2 parece tener una mejor estabilidad en la precisión en comparación con ResNet50v2. Ambas curvas para la clase normal muestran un excelente rendimiento en ambos modelos, con alta precisión y recall en ambos casos, indicando un rendimiento casi perfecto en la detección de imágenes normales

La precisión general del modelo fue del 94%, lo que indicó que el modelo tuvo un buen rendimiento global. Las métricas detalladas para cada clase mostraron que el modelo ResNet50v2 tuvo un rendimiento similar de clasificación a EfficientNetv2 tanto para las clases benignas como para las malignas y normales, con f1-scores elevados que sugirieron un balance adecuado entre precisión y recall.

La comparación entre los dos modelos mostró que, aunque ambos modelos tuvieron un alto rendimiento, EfficientNetV2 fue ligeramente superior en términos de precisión, recall y F1-score. EfficientNetV2 alcanzó una precisión del 95%, mientras que ResNet50v2 alcanzó un 94%. Esta diferencia, aunque pequeña, puede ser significativa en un entorno clínico donde cada porcentaje adicional en la precisión puede traducirse en mejores diagnósticos para los pacientes.

EfficientNetV2 mostró una mayor estabilidad y una menor tasa de predicciones incorrectas en la clase benigna, lo que es crucial para aplicaciones clínicas. La alta precisión en la detección de tumores benignos ayuda a reducir los falsos positivos, evitando así procedimientos médicos innecesarios y ansiedad para los pacientes.

Las matrices de confusión de ambos modelos resaltaron la capacidad de EfficientNetV2 para manejar mejor las clases de datos. EfficientNetV2 tuvo menos errores en la clasificación de imágenes benignas y malignas, mostrando una mayor eficacia en la diferenciación entre estas dos categorías críticas. ResNet50v2, aunque también efectivo, mostró más fluctuaciones y errores en las predicciones, lo que podría indicar una menor robustez en su capacidad de generalización.



**Ilustración 30.** Distribución de predicciones ResNet50v2 y EfficientNetv2

Fuente: Elaboración propia

### 5.6.1. VALIDACIÓN MÉDICA

Para fortalecer la validez clínica de los resultados obtenidos a partir del modelo de redes neuronales convolucionales EfficientNetV2, se llevó a cabo un proceso de validación médica. Este proceso consistió en la consulta con un radiólogo para revisar las imágenes de ecografías mamarias clasificadas por el modelo. El objetivo fue comprobar la precisión del diagnóstico probable proporcionado por el modelo con la opinión profesional de un especialista en radiología.

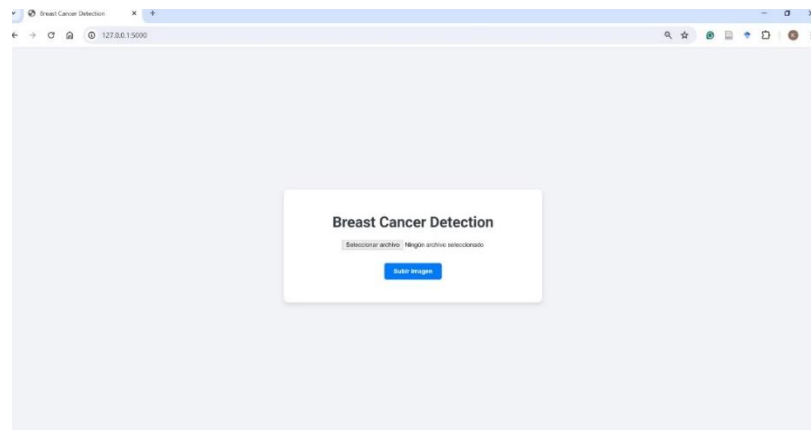
El primer paso fue la selección de imágenes. Se seleccionaron aleatoriamente un conjunto de imágenes de ecografías mamarias clasificadas como benignas, malignas y normales por el modelo. Se aseguró que el conjunto de imágenes incluyera casos correctos y aquellos con predicciones incorrectas para evaluar el rendimiento del modelo de manera integral.

Las imágenes seleccionadas se presentaron a un radiólogo con amplia experiencia en la interpretación de ecografías mamarias. El radiólogo revisó cada imagen de forma independiente, sin conocer la clasificación proporcionada por el modelo. Para cada imagen, el radiólogo emitió un diagnóstico basado en su experiencia profesional.

Posteriormente, se compararon los diagnósticos proporcionados por el radiólogo con las clasificaciones realizadas por el modelo de IA. Se identificaron las coincidencias y discrepancias entre las predicciones del modelo y la opinión del radiólogo. La validación médica proporciona una confirmación importante de que el modelo de IA, EfficientNetV2, puede ser una herramienta útil en la detección de tumores en imágenes de ecografías mamarias.

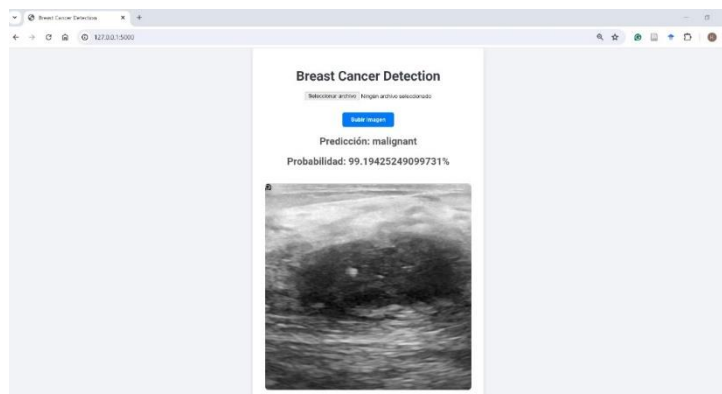
En una segunda visita al hospital, se evaluaron imágenes de pacientes ya confirmados con cáncer benigno y maligno. El modelo fue capaz de acertar en 8 diagnósticos correctos y cometió 1 error. Sin embargo, se presentaron algunas limitantes durante esta fase, ya que se utilizó una cámara de celular para exportar las fotos, debido a restricciones que impedían importar las imágenes directamente.

La interfaz de usuario empleada para la validación permitió subir archivos de imágenes y mostrar el resultado del diagnóstico probable con su respectiva probabilidad. A pesar de las limitaciones encontradas, el modelo obtuvo una media de precisión de 0.95635, lo cual es un indicador de su potencial eficacia en un entorno clínico. La interfaz, como se muestra en las imágenes, simplificó el proceso de uso, permitiendo al usuario subir la imagen de la ecografía y recibir una predicción inmediata del modelo.



**Ilustración 31.** Interfaz del Prototipo

Fuente: Elaboración propia



**Ilustración 32.** Interfaz del Prototipo

Fuente: Elaboración propia

### 5.7. REGISTRO SOBRE LA EFECTIVIDAD DEL MODELO

En la revisión de la literatura, se analizaron diversos estudios que emplearon una variedad de arquitecturas de entrenamiento y técnicas de inteligencia artificial para la detección del cáncer de mama. Estos estudios utilizaron desde Máquinas de Vectores de Soporte (SVM) y Redes Neuronales Convolucionales (CNN) hasta LightGBM y técnicas de fusión de imágenes y aprendizaje profundo con ajuste fino. Los conjuntos de datos variaron considerablemente, abarcando mamografías, imágenes histopatológicas y ecografías mamarias, lo que demuestra la versatilidad y aplicabilidad de estas tecnologías en diferentes contextos clínicos.

**Tabla 11.** Comparación de Resultados con Antecedentes

| Ítem | Título  | Arquitectura de Entrenamiento  | Resultado de la investigación | Conjunto de Datos Imágenes  |
|------|---|--|-------------------------------|---|
| 1    | An Automatic Detection of Breast Cancer Diagnosis and Prognosis Based on Machine Learning Using Ensemble of Classifiers | SVM (Máquinas de Vectores de Soporte)<br>LR (Regresión Logística)<br>NB (Naive Bayes)<br>DT (Árbol de Decisión)<br>ANN (Red Neuronal Artificial) | Valor de precisión del 98.33% | Wisconsin Breast Cancer (Diagnosis),<br>Wisconsin Breast Cancer (Prognosis) |

**Tabla 11. (Continuación)**

| Ítem | Título   | Arquitectura de Entrenamiento  | Resultados de investigación  | Conjunto de Datos Imágenes  |
|------|--|--|--|---|
| 2    | Artificial intelligence-based computer-assisted detection/diagnosis (AI-CAD) for screening mammography: Outcomes of AI-CAD in the mammographic interpretation workflow | Redes Neuronales Convolucionales Profundas (CNNs)  | Valor de Sensibilidad: 82.1%   | 6499 mamografías de 5228 mujeres  |
| 3    | LightGBM: A Leading Force in Breast Cancer Diagnosis Through Machine Learning and Image Processing   | LightGBM   | Precisión sin aumento: 95.70%<br>Incremento: 99.18% (400x)                   | BreaKHis, 7,909 imágenes histopatológicas de tejido tumoral de mama (40x, 100x, 200x, 400x)         |
| 4    | Development of digital breast tomosynthesis and diffuse optical tomography fusion imaging for breast cancer detection  | Técnica de imagen de fusión DBT/DOT con sistemas DBT y DOT multicanal                              | DBT: AUC 0.783<br>DOT: AUC 0.736<br>DBT/DOT: AUC 0.964                       | 28 muestras de tomógrafo  |
| 5    | Differentiation of breast lesions on dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) using deep transfer learning based on DenseNet201                  | DenseNet201  | Valor de precisión del 93.90%  | 8400 imágenes de MRI  |
| 6    | Prediction of Breast Cancer, Comparative Review of Machine Learning Techniques, and Their Analysis   | ANN, Logistic Regression, KNN, Decision Tree, Naive Bayes, SVM, Random Forest, Ensemble Techniques | SVM: 99.68%,<br>Random Forest: 92.2%,<br>Naive Bayes y KNN: 98%,<br>ANN: 97% | WBCD, WDBC, Bioinformatics and Medical Science Classification Dataset, MIAS Database, SEER Database |
| 7    | Artificial intelligence system reduces false-positive findings in the interpretation of breast ultrasound exams  | Deep Neural Network (DNN)  | AUROC: 0.976   | NYU Breast Ultrasound Dataset<br>Breast Ultrasound Images (BUSI)                                    |
| 8    | Presente investigación   | EfficientNetv2   | AUROC: 0.98,<br>Precisión: 95%   | Base de datos: 780 imágenes<br>Aumentación: 1877 imágenes   |

Fuente: Elaboración propia

En la investigación se evaluaron diferentes modelos de inteligencia artificial para la detección del cáncer de mama, logrando precisiones significativas con distintas arquitecturas, se realizó una comparación de los resultados obtenidos con EfficientNetV2 con los resultados de estudios previos en el campo. Esta comparación tuvo como objetivo destacar las fortalezas y mejoras del modelo.

## **5.8. LIMITACIONES**

Durante el desarrollo de este proyecto, se identificaron varias limitaciones que podrían haber afectado los resultados y la eficacia del modelo de red neuronal para la detección de tumores en imágenes de ecografías mamarias. La base de datos utilizada no era lo suficientemente extensa, limitándose a un conjunto específico de 780 imágenes de ecografía por ultrasonido. Esta restricción en el tamaño del conjunto de datos puede afectar la capacidad de la red neuronal para generalizar los datos.

La base de datos utilizada contenía imágenes de tres categorías (normales, benignas y malignas), pero no describe una variedad de subtipos de lesiones. La capacidad del modelo para diferenciar entre estos diferentes subtipos de tumores mamarios podría estar limitada, afectando diagnósticos más complejos.

Adicionalmente, el entrenamiento de la red neuronal se realizó únicamente utilizando el CPU de la computadora, sin la inclusión de una GPU, lo que dio como resultado un tiempo de entrenamiento más prolongado. La falta de una GPU adecuada limitó la implementación de librerías que podrían haber mejorado el rendimiento diagnóstico.

## **5.9. ÁREAS DE MEJORA**

En esta sección, se detallan varias recomendaciones para mejorar y optimizar el proyecto de detección de tumores en imágenes de ecografías mamarias utilizando inteligencia artificial. Estas recomendaciones buscan abordar las limitaciones identificadas y potenciar la eficacia y aplicabilidad del modelo en diferentes contextos clínicos y geográficos.

- Se debe recopilar una base de datos más amplia y diversa de imágenes de ecografía mamarias. Recopilar datos locales, incluyendo imágenes y datos de pacientes de diversas regiones de Honduras, esto asegurará que el modelo de IA sea representativo y efectivo para la población local. La inclusión de datos locales no solo mejorará la precisión del modelo en el contexto específico del país, sino que también contribuirá a una mejor comprensión de las características particulares del cáncer de mama en la población hondureña.

- Utilizar una GPU (Unidad de Procesamiento Gráfico) para el entrenamiento del modelo puede acelerar significativamente el proceso. No solo reducirá el tiempo de entrenamiento, sino que también permitirá experimentar con arquitecturas más sofisticadas, mejorando potencialmente la precisión y eficiencia del modelo.

- Desarrollar e implementar el programa en una plataforma basada en la nube para que pueda ser accesible por múltiples médicos simultáneamente, sin la necesidad de hardware especializado en cada ubicación. Esta infraestructura permitirá que el modelo sea utilizado en tiempo real por profesionales de la salud en diversas localizaciones, facilitando la colaboración y la segunda opinión entre médicos.

- Adaptar y optimizar el programa para su uso en áreas rurales y de bajo recurso, donde el acceso a especialistas en diagnóstico puede ser limitado. Esto incluye la creación de una interfaz sencilla y robusta que pueda ser utilizada con equipos de ecografía básicos y en condiciones de conectividad limitada.

## VI. CONCLUSIONES

- Del proceso investigativo efectuado, se obtuvo una evaluación sistemática de seis variables en tres diferentes modelos de inteligencia artificial (IA) aplicados con el propósito de la clasificación de tumores empleando imágenes de ecografía mamaria. El mayor aporte técnico de esta investigación se manifestó en la variable de precisión con el modelo EfficientNetV2. Con la aplicación de este modelo, se obtuvo un valor del 95%. Dada la naturaleza exploratoria de la aplicación de este modelo de IA en imágenes de ecografía mamaria; se identificó que podría tener potencial como alternativa de apoyo al diagnóstico médico al registrarse un valor aceptable de precisión para tumores benignos y malignos.
- Se seleccionó una base de datos diversificada, cumpliendo con los parámetros buscados en la segregación de las imágenes de ecografías mamarias en las tres categorías deseadas. Esta base de datos fue revisada por el Médico Mohamed Hamed en el Hospital Baheya. Posteriormente esto permitió la aplicación de las técnicas de preprocesamiento y aumentación de datos estipuladas, robusteciendo la base de datos para el entrenamiento del modelo.
- Se realizó el análisis de la influencia del tamaño del lote en la precisión y pérdida del modelo durante su entrenamiento para identificar el tamaño de lote óptimo, luego se encontró que tamaños de lote más pequeños tiende a proporcionar una mayor precisión y estabilidad durante el entrenamiento, mientras que tamaños de lote más grandes introdujeron fluctuaciones y afectaron negativamente el rendimiento del modelo. Este hallazgo contribuye a un entrenamiento más eficiente realizándolo con un tamaño de lote pequeño, tomando en cuenta que se aumenta el tiempo de entrenamiento.
- Se hizo una comparación entre diferentes arquitecturas de redes neuronales convolucionales, como EfficientNetV2, ResNet50 y VGG16, por lo tanto, se reveló que EfficientNetV2 fue el modelo más eficiente y preciso para la clasificación de tumores mamarios, luego de que tuviera un mayor rendimiento con el conjunto de datos de entrenamiento, evaluando métricas como la precisión, la pérdida y el tiempo de

entrenamiento. EfficientNetV2 demostró un rendimiento superior en términos de precisión diagnóstica.

- Se logró la validación del modelo de inteligencia artificial de EfficientNetv2 utilizando un conjunto de datos independiente, el modelo confirmó su capacidad de generalización y efectividad con datos nuevos. Al analizar métricas como la sensibilidad, precisión y F1-score, se comprobó que el modelo adaptado es capaz de proporcionar diagnósticos, asegurando como alternativa para su uso en el diagnóstico de clasificación de tipos de tumores.

## VII. REFERENCIAS BIBLIOGRÁFICAS

- Ahn, J. S., Shin, S., Yang, S.-A., Park, E. K., Kim, K. H., Cho, S. I., Ock, C.-Y., & Kim, S. (2023). Artificial Intelligence in Breast Cancer Diagnosis and Personalized Medicine. *Journal of Breast Cancer*, 26(5), 405-435. <https://doi.org/10.4048/jbc.2023.26.e45>
- Al-Dhabyani, W., Gomaa, M., Khaled, H., & Fahmy, A. (2020). Dataset of breast ultrasound images. *Data in Brief*, 28, 104863. <https://doi.org/10.1016/j.dib.2019.104863>
- Arshad, W., Masood, T., Mahmood, T., Jaffar, A., Alamri, F. S., Bahaj, S. A. O., & Khan, A. R. (2023). Cancer Unveiled: A Deep Dive Into Breast Tumor Detection Using Cutting-Edge Deep Learning Models. *IEEE Access*, 11, 133804-133824. <https://doi.org/10.1109/ACCESS.2023.3335604>
- Chae, E. Y., Kim, H. H., Sabir, S., Kim, Y., Kim, H., Yoon, S., Ye, J. C., Cho, S., Heo, D., Kim, K. H., Bae, Y. M., & Choi, Y.-W. (2020). Development of digital breast tomosynthesis and diffuse optical tomography fusion imaging for breast cancer detection. *Scientific Reports*, 10(1), 13127. <https://doi.org/10.1038/s41598-020-70103-0>
- Dileep, G., & Gianchandani Gyani, S. G. (s. f.). Artificial Intelligence in Breast Cancer Screening and Diagnosis. *Cureus*, 14(10), e30318. <https://doi.org/10.7759/cureus.30318>
- Dileep, G., Gyani, S. G. G., Dileep, G., & Gyani, S. G. G. (2022). Artificial Intelligence in Breast Cancer Screening and Diagnosis. *Cureus*, 14(10). <https://doi.org/10.7759/cureus.30318>
- Ecografía mamaria por ultrasonidos. Clínica Universidad de Navarra.* (s. f.). <https://www.cun.es>. Recuperado 30 de junio de 2024, de <https://www.cun.es/enfermedades-tratamientos/pruebas-diagnosticas/ecografia-mamaria>

Freeman, K., Geppert, J., Stinton, C., Todkill, D., Johnson, S., Clarke, A., & Taylor-Phillips, S. (2021). Use of artificial intelligence for image analysis in breast cancer screening programmes: Systematic review of test accuracy. *BMJ*, *374*, n1872. <https://doi.org/10.1136/bmj.n1872>

Hanna, T. P., King, W. D., Thibodeau, S., Jalink, M., Paulin, G. A., Harvey-Jones, E., O'Sullivan, D. E., Booth, C. M., Sullivan, R., & Aggarwal, A. (2020). Mortality due to cancer treatment delay: Systematic review and meta-analysis. *BMJ*, *371*, m4087. <https://doi.org/10.1136/bmj.m4087>

Huang, Z., & Chen, D. (2022). A Breast Cancer Diagnosis Method Based on VIM Feature Selection and Hierarchical Clustering Random Forest Algorithm. *IEEE Access*, *10*, 3284-3293. <https://doi.org/10.1109/ACCESS.2021.3139595>

IBM. (2021, octubre 6). <https://www.ibm.com/mx-es/topics/artificial-intelligence>

INSTITUTO NACIONAL DEL CANCER (nciglobal,ncienterprise). (2011, febrero 2). [nciAppModulePage]. <https://www.cancer.gov/espanol/publicaciones/diccionarios/diccionario-cancer/def/cancer-de-mama>

Kanber, B. M., Smadi, A. A., Noaman, N. F., Liu, B., Gou, S., & Alsmadi, M. K. (2024). LightGBM: A Leading Force in Breast Cancer Diagnosis Through Machine Learning and Image Processing. *IEEE Access*, *12*, 39811-39832. <https://doi.org/10.1109/ACCESS.2024.3375755>

Kausar, T., Lu, Y., & Kausar, A. (2023). Breast Cancer Diagnosis Using Lightweight Deep Convolution Neural Network Model. *IEEE Access*, *11*, 124869-124886. <https://doi.org/10.1109/ACCESS.2023.3326478>

- Kim, C.-M., Park, R. C., & Hong, E. J. (2020). Breast Mass Classification Using eLFA Algorithm Based on CRNN Deep Learning Model. *IEEE Access*, *8*, 197312-197323.  
<https://doi.org/10.1109/ACCESS.2020.3034914>
- Mao, X., He, W., Humphreys, K., Eriksson, M., Holowko, N., Yang, H., Tapia, J., Hall, P., & Czene, K. (2024). Breast Cancer Incidence After a False-Positive Mammography Result. *JAMA Oncology*, *10*(1), 63-70. <https://doi.org/10.1001/jamaoncol.2023.4519>
- Meng, M., Zhang, M., Shen, D., & He, G. (2022). Differentiation of breast lesions on dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) using deep transfer learning based on DenseNet201. *Medicine*, *101*(45), e31214.  
<https://doi.org/10.1097/MD.00000000000031214>
- Naseem, U., Rashid, J., Ali, L., Kim, J., Haq, Q. E. U., Awan, M. J., & Imran, M. (2022). An Automatic Detection of Breast Cancer Diagnosis and Prognosis Based on Machine Learning Using Ensemble of Classifiers. *IEEE Access*, *10*, 78242-78252.  
<https://doi.org/10.1109/ACCESS.2022.3174599>
- OMS. (s. f.). *Promoting cancer early diagnosis*. Recuperado 21 de abril de 2024, de <https://www.who.int/activities/promoting-cancer-early-diagnosis>
- Osman, A. H., & Aljahdali, H. M. A. (2020). An Effective of Ensemble Boosting Learning Method for Breast Cancer Virtual Screening Using Neural Network Model. *IEEE Access*, *8*, 39165-39174. <https://doi.org/10.1109/ACCESS.2020.2976149>
- ¿Qué es el sobreajuste? | IBM. (2024, abril 3). <https://www.ibm.com/es-es/topics/overfitting>

¿Qué son las redes neuronales convolucionales? | 3 cosas que debe saber. (s. f.-a). Recuperado 30 de junio de 2024, de <https://la.mathworks.com/discovery/convolutional-neural-network.html>

RAE. (2024, enero 19). *Tumor* | *Diccionario del estudiante*. «Diccionario del estudiante». <https://www.rae.es/diccionario-estudiante/tumor>

Saber, A., Sakr, M., Abo-Seida, O. M., Keshk, A., & Chen, H. (2021). A Novel Deep-Learning Model for Automatic Detection and Classification of Breast Cancer Using the Transfer-Learning Technique. *IEEE Access*, 9, 71194-71209. <https://doi.org/10.1109/ACCESS.2021.3079204>

Shen, Y., Shamout, F. E., Oliver, J. R., Witowski, J., Kannan, K., Park, J., Wu, N., Huddleston, C., Wolfson, S., Millet, A., Ehrenpreis, R., Awal, D., Tyma, C., Samreen, N., Gao, Y., Chhor, C., Gandhi, S., Lee, C., Kumari-Subaiya, S., ... Geras, K. J. (2021). Artificial intelligence system reduces false-positive findings in the interpretation of breast ultrasound exams. *Nature Communications*, 12(1), 5645. <https://doi.org/10.1038/s41467-021-26023-2>

*Transfer Learning*. (s. f.-b). Recuperado 30 de junio de 2024, de <https://la.mathworks.com/discovery/transfer-learning.html>

Wang, Z., Li, M., Wang, H., Jiang, H., Yao, Y., Zhang, H., & Xin, J. (2019). Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion With CNN Deep Features. *IEEE Access*, 7, 105146-105158. <https://doi.org/10.1109/ACCESS.2019.2892795>

Yoon, J. H., Han, K., Suh, H. J., Youk, J. H., Lee, S. E., & Kim, E.-K. (2023). Artificial intelligence-based computer-assisted detection/diagnosis (AI-CAD) for screening mammography: Outcomes of AI-CAD in the mammographic interpretation workflow. *European Journal of Radiology Open*, 11. <https://doi.org/10.1016/j.ejro.2023.100509>

- Zhang, S., Yang, B., Yang, H., Zhao, J., Zhang, Y., Gao, Y., Monteiro, O., Zhang, K., Liu, B., & Wang, S. (2024). Potential rapid intraoperative cancer diagnosis using dynamic full-field optical coherence tomography and deep learning: A prospective cohort study in breast cancer patients. *Science Bulletin*. <https://doi.org/10.1016/j.scib.2024.03.061>
- Sampieri, R. H., Collado, C. F., Lucio, P. B., Valencia, S. M., & Torres, C. P. M. (2014). Metodología de la investigación.
- De Carvalho, O. L. F., De Carvalho, O. A., Junior, De Albuquerque, A. O., Orlandi, A. G., Hirata, I., Borges, D. L., Gomes, R. A. T., & Guimarães, R. F. (2023). A Data-Centric Approach for Wind Plant Instance-Level Segmentation Using Semantic Segmentation and GIS. *Remote Sensing*, 15(5), 1240. <https://doi.org/10.3390/rs15051240>
- Al-Dhabyani W, Gomaa M, Khaled H, Fahmy A. Dataset of breast ultrasound images. *Data in Brief*. 2020 Feb;28:104863. DOI: 10.1016/j.dib.2019.104863.
- Cáncer de mama. (2024, 1 mayo). OPS/OMS | Organización Panamericana de la Salud. <https://www.paho.org/es/temas/cancer>.

# ANEXOS

```
Epoch 28/300
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m8s-[0m 1s/step - accuracy: 0.9854 - loss: 0.0432
Epoch 28: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m298s-[0m 1s/step - accuracy: 0.9854 - loss: 0.0432 - val_accuracy: 0.9351 - val_loss: 0.2777 - learning_rate: 1.0000e-04
Epoch 29/300
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m8s-[0m 1s/step - accuracy: 0.9794 - loss: 0.0639
Epoch 29: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m312s-[0m 2s/step - accuracy: 0.9794 - loss: 0.0639 - val_accuracy: 0.9351 - val_loss: 0.4084 - learning_rate: 1.0000e-04
Epoch 30/300
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m8s-[0m 2s/step - accuracy: 0.9793 - loss: 0.0676
Epoch 30: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m382s-[0m 2s/step - accuracy: 0.9793 - loss: 0.0675 - val_accuracy: 0.9221 - val_loss: 0.4561 - learning_rate: 1.0000e-04
Epoch 31/300
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m0s-[0m 1s/step - accuracy: 0.9816 - loss: 0.0666
Epoch 31: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m311s-[0m 2s/step - accuracy: 0.9816 - loss: 0.0667 - val_accuracy: 0.8961 - val_loss: 0.4023 - learning_rate: 1.0000e-04
Epoch 32/300
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m8s-[0m 1s/step - accuracy: 0.9862 - loss: 0.0313
Epoch 32: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m294s-[0m 1s/step - accuracy: 0.9862 - loss: 0.0313 - val_accuracy: 0.9091 - val_loss: 0.4660 - learning_rate: 1.0000e-04
Epoch 33/300
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m8s-[0m 2s/step - accuracy: 0.9845 - loss: 0.0498
Epoch 33: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m334s-[0m 2s/step - accuracy: 0.9845 - loss: 0.0499 - val_accuracy: 0.8831 - val_loss: 0.5613 - learning_rate: 1.0000e-04
Epoch 34/300
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m8s-[0m 2s/step - accuracy: 0.9732 - loss: 0.0793
Epoch 34: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m359s-[0m 2s/step - accuracy: 0.9732 - loss: 0.0793 - val_accuracy: 0.9091 - val_loss: 0.4753 - learning_rate: 1.0000e-04
Epoch 35/300
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m8s-[0m 1s/step - accuracy: 0.9829 - loss: 0.0463
Epoch 35: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m308s-[0m 1s/step - accuracy: 0.9829 - loss: 0.0462 - val_accuracy: 0.9481 - val_loss: 0.2461 - learning_rate: 1.0000e-04
Epoch 36/300
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m0s-[0m 1s/step - accuracy: 0.9789 - loss: 0.0708
Epoch 36: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m294s-[0m 1s/step - accuracy: 0.9789 - loss: 0.0708 - val_accuracy: 0.9221 - val_loss: 0.3227 - learning_rate: 1.0000e-04
Epoch 37/300
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m8s-[0m 2s/step - accuracy: 0.9900 - loss: 0.0329
Epoch 37: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m354s-[0m 2s/step - accuracy: 0.9900 - loss: 0.0330 - val_accuracy: 0.8571 - val_loss: 0.5436 - learning_rate: 1.0000e-04
Epoch 38/300
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m8s-[0m 2s/step - accuracy: 0.9931 - loss: 0.0283
Epoch 38: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m345s-[0m 2s/step - accuracy: 0.9931 - loss: 0.0284 - val_accuracy: 0.8961 - val_loss: 0.5071 - learning_rate: 1.0000e-04
Epoch 39/300
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m8s-[0m 1s/step - accuracy: 0.9886 - loss: 0.0322
Epoch 39: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m306s-[0m 1s/step - accuracy: 0.9886 - loss: 0.0322 - val_accuracy: 0.9221 - val_loss: 0.3121 - learning_rate: 1.0000e-04
Epoch 40/300
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m8s-[0m 1s/step - accuracy: 0.9900 - loss: 0.0269
Epoch 40: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m294s-[0m 1s/step - accuracy: 0.9900 - loss: 0.0269 - val_accuracy: 0.8831 - val_loss: 0.5367 - learning_rate: 1.0000e-04
Epoch 41/300
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m8s-[0m 2s/step - accuracy: 0.9904 - loss: 0.0222
Epoch 41: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m369s-[0m 2s/step - accuracy: 0.9904 - loss: 0.0222 - val_accuracy: 0.8961 - val_loss: 0.4435 - learning_rate: 1.0000e-04
Epoch 42/300
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m8s-[0m 74s/step - accuracy: 0.9867 - loss: 0.0349
Epoch 42: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m1517s-[0m 74s/step - accuracy: 0.9867 - loss: 0.0349 - val_accuracy: 0.9091 - val_loss: 0.5059 - learning_rate: 1.0000e-04
Epoch 43/300
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m8s-[0m 2s/step - accuracy: 0.9881 - loss: 0.0286
Epoch 43: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m286/286](0m +32s) [0m]{37m}[0m +[1m8s-[0m 2s/step - accuracy: 0.9881 - loss: 0.0286
Epoch 43: ReduceLRonPlateau reducing learning rate to 9.99999747378752e-06.
```

Ilustración 33. Épocas tamaño de lote 8 EfficientNetv2

Fuente: Elaboración propia

```
[1m274/274](0m +32s) [0m]{37m}[0m +[1m8s-[0m 982ms/step - accuracy: 0.9629 - loss: 0.1179
Epoch 13: val_accuracy did not improve from 0.92288
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m274/274](0m +32s) [0m]{37m}[0m +[1m272s-[0m 991ms/step - accuracy: 0.9629 - loss: 0.1179 - val_accuracy: 0.9091 - val_loss: 0.2853 - learning_rate: 1.0000e-04
Epoch 14/100
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m274/274](0m +32s) [0m]{37m}[0m +[1m8s-[0m 979ms/step - accuracy: 0.9690 - loss: 0.1012
Epoch 14: val_accuracy did not improve from 0.92288
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m274/274](0m +32s) [0m]{37m}[0m +[1m271s-[0m 988ms/step - accuracy: 0.9690 - loss: 0.1011 - val_accuracy: 0.9221 - val_loss: 0.2938 - learning_rate: 1.0000e-04
Epoch 15/100
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m274/274](0m +32s) [0m]{37m}[0m +[1m8s-[0m 421s/step - accuracy: 0.9744 - loss: 0.0750
Epoch 15: val_accuracy improved from 0.92288 to 0.93506, saving model to C:/Users/LAGOS/Desktop/articulos/Breast Cancer/32epochEfficientNetV2.keras
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m274/274](0m +32s) [0m]{37m}[0m +[1m1343s-[0m 421s/step - accuracy: 0.9744 - loss: 0.0751 - val_accuracy: 0.9351 - val_loss: 0.3170 - learning_rate: 1.0000e-04
Epoch 16/100
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m274/274](0m +32s) [0m]{37m}[0m +[1m8s-[0m 979ms/step - accuracy: 0.9783 - loss: 0.0823
Epoch 16: val_accuracy improved from 0.93506 to 0.94805, saving model to C:/Users/LAGOS/Desktop/articulos/Breast Cancer/32epochEfficientNetV2.keras
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m274/274](0m +32s) [0m]{37m}[0m +[1m273s-[0m 993ms/step - accuracy: 0.9782 - loss: 0.0823 - val_accuracy: 0.9481 - val_loss: 0.2524 - learning_rate: 1.0000e-04
Epoch 17/100
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m274/274](0m +32s) [0m]{37m}[0m +[1m8s-[0m 979ms/step - accuracy: 0.9731 - loss: 0.0688
Epoch 17: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m274/274](0m +32s) [0m]{37m}[0m +[1m271s-[0m 987ms/step - accuracy: 0.9731 - loss: 0.0688 - val_accuracy: 0.8182 - val_loss: 0.5793 - learning_rate: 1.0000e-04
Epoch 18/100
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m274/274](0m +32s) [0m]{37m}[0m +[1m8s-[0m 979ms/step - accuracy: 0.9783 - loss: 0.0823
Epoch 18: val_accuracy did not improve from 0.94805
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
[1m274/274](0m +32s) [0m]{37m}[0m +[1m294s-[0m 1s/step - accuracy: 0.9749 - loss: 0.0604 - val_accuracy: 0.9221 - val_loss: 0.3078 - learning_rate: 1.0000e-04
Epoch 41: early stopping
Mejor precisión de validación: 0.948051929473877
C:\Users\LAGOS\AppData\Local\Programs\Python\Python312\python.exe: UserWarning: Starting a Matplotlib GUI outside of the main thread will likely fail.
plt.plot(result.history['accuracy'], label='precisión de entrenamiento')
C:\Users\LAGOS\AppData\Local\Programs\Python\Python312\python.exe: UserWarning: Starting a Matplotlib GUI outside of the main thread will likely fail.
plt.show()
```

Ilustración 34. Épocas tamaño de lote 6 EfficientNetv2

Fuente: Elaboración propia

```

ss: 0.0440
Epoch 36: val_accuracy did not improve from 0.93506

Epoch 36: ReduceLROnPlateau reducing learning rate to 9.999999747378752e-06.
[Progress bar]
[1m274/274][0m ][32m
[1m1262s][0m 5s/step - accuracy: 0.9858 - loss: 0.0440 - val_accuracy: 0.8831 - val_loss: 0.4680 - learning_rate: 1.0000e-04
Epoch 36: early stopping
Mejor precisión de validación: 0.9350649118423462

```

**Ilustración 35.** Épocas tamaño de lote 12 EfficientNetv2

Fuente: Elaboración propia

```

Epoch 42: val_accuracy did not improve from 0.93506

Epoch 42: ReduceLROnPlateau reducing learning rate to 9.999999747378752e-06.
[Progress bar]
[1m52/52][0m ][32m
270s][0m 5s/step - accuracy: 0.9911 - loss: 0.0254 - val_accuracy: 0.8442 - val_loss: 0.5213 - learning_rate: 1.0000e-04
Epoch 42: early stopping
Mejor precisión de validación: 0.9090909361839294

```

**Ilustración 36.** Épocas tamaño de lote 32 EfficientNetv2

Fuente: Elaboración propia

```

Epoch 60: val_accuracy improved from 0.92208 to 0.93506, saving model to C:/Users/LAGOS/Desktop/articulos/Breast Cancer/6batchResNet50V2.keras
[Progress bar]
[1m274/274][0m ][32m
[1m1166s][0m 4s/step - accuracy: 0.9747 - loss: 0.0767 - val_accuracy: 0.9351 - val_loss: 0.3322 - learning_rate: 1.0000e-04
Epoch 61/100
[1m 1/274][0m ][37m
Epoch 85: val_accuracy did not improve from 0.93506

Epoch 85: ReduceLROnPlateau reducing learning rate to 9.999999747378752e-06.
[Progress bar]
[1m274/274][0m ][32m
[1m1139s][0m 4s/step - accuracy: 0.9846 - loss: 0.0449 - val_accuracy: 0.8961 - val_loss: 0.4251 - learning_rate: 1.0000e-04
Epoch 85: early stopping
Mejor precisión de validación: 0.9350649118423462

```

**Ilustración 37.** Épocas tamaño de lote 6 ResNet50v2

Fuente: Elaboración propia

| Layer (type)                                      | Output Shape       | Param #    |
|---|--------------------|------------|
| resnet50v2 (Functional)                           | (None, 7, 7, 2048) | 23,564,800 |
| global_average_pooling2d (GlobalAveragePooling2D) | (None, 2048)       | 0          |
| dense (Dense)                                     | (None, 1024)       | 2,098,176  |
| dropout (Dropout)                                 | (None, 1024)       | 0          |
| dense_1 (Dense)                                   | (None, 3)          | 3,075      |

Total params: 76,907,275 (293.38 MB)  
Trainable params: 25,620,611 (97.73 MB)  
Non-trainable params: 45,440 (177.50 KB)  
Optimizer params: 51,241,224 (195.47 MB)

**Ilustración 38.** Arquitectura ResNet50v2

Fuente: Elaboración propia

| Layer (type)   | Output Shape       | Param #    |
|--|--------------------|------------|
| efficientnetv2-s (Functional)                        | (None, 7, 7, 1280) | 20,331,360 |
| global_average_pooling2d<br>(GlobalAveragePooling2D) | (None, 1280)       | 0          |
| dense (Dense)  | (None, 1024)       | 1,311,744  |
| dropout (Dropout)                                    | (None, 1024)       | 0          |
| dense_1 (Dense)                                      | (None, 3)          | 3,075      |

Total params: 64,630,795 (246.55 MB)  
Trainable params: 21,492,307 (81.99 MB)  
Non-trainable params: 153,872 (601.06 KB)  
Optimizer params: 42,984,616 (163.97 MB)

**Ilustración 39.** Arquitectura EfficientNetv2

Fuente: Elaboración propia

