



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Dirección General de Estudios de Posgrado
Facultad de Ingeniería de Sistemas e Informática
Unidad de Posgrado

**Modelo Basado en Aprendizaje Profundo para
Clasificar Células Neoplásicas en Imágenes de
Citología de Líquido Pleural**

TESIS

Para optar el Grado Académico de Doctora en Ingeniería de
Sistemas e Informática

AUTOR

Frida Mereyda LÓPEZ CÓRDOVA

ASESOR

Dr. Hugo Froilán VEGA HUERTA

Lima, Perú

2025



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

López, F. (2025). *Modelo Basado en Aprendizaje Profundo para Clasificar Células Neoplásicas en Imágenes de Citología de Líquido Pleural*. [Tesis de doctorado, Universidad Nacional Mayor de San Marcos, Facultad de Ingeniería de Sistemas e Informática, Unidad de Posgrado]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Frida Mereyda López Córdova
Tipo de documento de identidad	DNI
Número de documento de identidad	06799158
URL de ORCID	https://orcid.org/0000-0002-0494-6800
Datos de asesor	
Nombres y apellidos	Hugo Froilán Vega Huerta
Tipo de documento de identidad	DNI
Número de documento de identidad	06147737
URL de ORCID	https://orcid.org/0000-0002-4268-5808
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Ciro Rodríguez Rodríguez
Tipo de documento	DNI
Número de documento de identidad	06040421
Miembro del jurado 1	
Nombres y apellidos	Nora Bertha la Serna Palomino
Tipo de documento	DNI
Número de documento de identidad	07665297
Miembro del jurado 2	
Nombres y apellidos	Pedro Martín Lezama Gonzales

Tipo de documento	DNI
Número de documento de identidad	09656793
Datos de investigación	
Línea de investigación	C.0.3.3. Desarrollo de modelos y aplicaciones de las tecnologías de información y comunicaciones
Grupo de investigación	Innovando Sistemas Inteligentes
Agencia de financiamiento	Sin financiamiento
Ubicación geográfica de la investigación	Universidad Nacional Mayor de San Marcos País: Perú Departamento: Lima Provincia: Lima Distrito: Lima Calle: Av. Universitaria cruce con Av. Venezuela cuadra 34 Latitud: -12.058333° Longitud: -77.083333°
Año o rango de años en que se realizó la investigación	Marzo 2022 – Diciembre 2024
URL de disciplinas OCDE	Ingeniería de sistemas y comunicaciones https://purl.org/pe-repo/ocde/ford#2.02.04 Ciencias de la computación https://purl.org/pe-repo/ocde/ford#1.02.00



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
Universidad del Perú. Decana de América
Facultad de Ingeniería de Sistemas e Informática
Vicedecanato de Investigación y Posgrado
Unidad de Posgrado

**ACTA DE SUSTENTACIÓN DE TESIS PARA OPTAR EL GRADO ACADÉMICO DE DOCTOR
EN INGENIERÍA DE SISTEMAS E INFORMÁTICA**

A los veinticuatro (24) días del mes de octubre de 2025, siendo las 10:00 a.m., se reunieron en el Auditorio Alfredo Celso Alva Bravo, el Jurado de Tesis conformado por los siguientes docentes:

Dr. Ciro Rodriguez Rodriguez (Presidente)
Dra. Nora Bertha La Serna Palomino (Miembro)
Dr. Pedro Martin Lezama Gonzales (Miembro)
Dr. Hugo Froilán Vega Huerta (Miembro Asesor)

Se inició la Sustentación invitando a la candidata a Doctora **FRIDA MEREYDA LÓPEZ CÓRDOVA**, para que realice la exposición oral de la tesis para optar el Grado Académico de Doctora en Ingeniería de Sistemas e Informática, siendo la Tesis intitulada:

**“Modelo Basado en Aprendizaje Profundo para
Clasificar Células Neoplásicas en Imágenes de Citología de Líquido Pleural”**

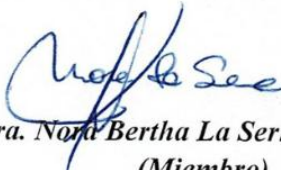
Concluida la exposición, los miembros del Jurado de Tesis procedieron a formular sus preguntas que fueron absueltas por el graduando; acto seguido se procedió a la evaluación correspondiente, habiendo obtenido la siguiente calificación:

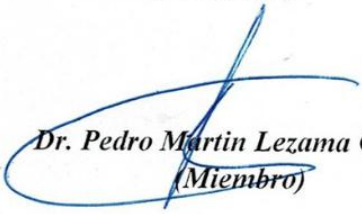
..... *Dieciséis* (16) *Aprobado con Honor*

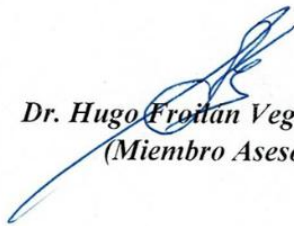
Por tanto, el presidente del Jurado, de acuerdo con el Reglamento General de Estudios de Posgrado, otorga a la Magister **FRIDA MEREYDA LÓPEZ CÓRDOVA** el Grado Académico de Doctora en Ingeniería de Sistemas e Informática.

Siendo las *11:20* horas, el presidente del Jurado de Tesis, da por concluido el acto académico de Sustentación de Tesis.


Dr. Ciro Rodriguez Rodriguez
(Presidente)


Dra. Nora Bertha La Serna Palomino
(Miembro)


Dr. Pedro Martin Lezama Gonzales
(Miembro)


Dr. Hugo Froilán Vega Huerta
(Miembro Asesor)



UNIVERSIDAD NACIONAL MAYOR DE

SAN MARCOS

Universidad del Perú, Decana de América

CERTIFICADO DE SIMILITUD

Yo, Hugo Froilán Vega Huerta en mi condición de asesor acreditado con Dictamen N° 000076-2024-UPG-VDIP-FISI/UNMSM de la tesis, cuyo título es **“MODELO BASADO EN APRENDIZAJE PROFUNDO PARA CLASIFICAR CÉLULAS NEOPLÁSICAS EN IMÁGENES DE CITOLOGÍA DE LÍQUIDO PLEURAL”** presentado por la magíster FRIDA MEREYDA LÓPEZ CÓRDOVA para optar el grado de Doctor en Ingeniería de Sistemas e Informática,

CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante la plataforma PIDS, el documento evaluado cuenta con el 14% de similitud, nivel PERMITIDO para continuar con los trámites correspondientes y para su publicación en el repositorio institucional.

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del grado/ título/ especialidad correspondiente.

Lima, 10 de abril de 2025

Firma del Asesor

Nombres y Apellidos del asesor: Hugo Froilán Vega Huerta

DNI: 06147737

DEDICATORIA

Hoy, al alcanzar este sueño, miro atrás y veo un camino lleno de esfuerzo, temor, sacrificio y amor. Este logro también es de quienes estuvieron a mi lado, apoyándome en cada momento difícil que tuve que superar.

A mi esposo, por su paciencia infinita, su amor inquebrantable y su fe en mí incluso en los momentos en que dudé de mí misma. A mis hijos, quienes han sido mi mayor inspiración, recordándome cada día el verdadero significado de la perseverancia y el amor incondicional. Verlos crecer mientras yo también crecía académicamente ha sido el motor que me impulsó a seguir adelante.

Ahora comprendo perfectamente lo que una vez leí y se grabó en mi mente: Eres más valiente de lo que crees, más inteligente de lo que piensas y más fuerte de lo que pareces.

Este doctorado es el fruto de noches largas, de retos superados y de un compromiso inquebrantable con mis sueños y mi familia. A ellos les dedico este logro, con la esperanza de que sepan que, con pasión y esfuerzo, todo es posible.

AGRADECIMIENTO

A Dios por la oportunidad de haber culminado mis estudios doctorales.

Siempre a mi madre, que desde el cielo me guía.

A mis profesores del programa doctoral, por todo el conocimiento que transmitieron y me motivaron a elaborar esta investigación.

Al profesor Dr Hugo Vega, por su paciencia, consejos, orientación y revisión de mi investigación.

A los profesores Dr. Augusto Bernuy, Dr. Yván Túpac, Dr. Ciro Rodríguez, Dr. Luis Rivera y Dra. Gisella Maquen por sus revisiones y recomendaciones en los artículos producto de esta tesis.

A los médicos especialistas Dra. Katherine Amaro y Dr. Jaime Cáceres por su asesoría en la especialidad de Anatomía Patológica.

A mis compañeros por sus observaciones y motivación para la presentación de mi investigación.

Al jurado revisor por su revisión y recomendación para mejorar este trabajo.

A Jimmy, mi amado esposo y compañero de vida, a mis hijos Elizabeth, Eduardo y Diego, que siempre me apoyaron y ayudaron a culminar mis estudios y tesis doctoral.

INDICE GENERAL

DEDICATORIA	ii
INDICE GENERAL.....	iv
LISTA DE FIGURAS	ix
LISTA DE TABLAS.....	xii
RESUMEN	xiv
ABSTRACT.....	xvi
CAPÍTULO 1. INTRODUCCIÓN.....	1
1.1. Antecedentes	3
1.2. Situación problemática	7
1.2.1. Importancia	9
1.2.2. Novedad	9
1.2.3. Interés	10
1.2.4. Viabilidad.....	10
1.3. Formulación del Problema.....	11
1.3.1. Problema principal	11

1.3.2. Problemas secundarios	11
1.4. Justificación	11
1.4.1. Justificación Teórica.....	13
1.4.2. Justificación Práctica	13
1.4.3. Motivación del estudio	13
1.5. Objetivos de la Investigación.....	14
1.5.1. Objetivo General	14
1.5.2. Objetivos Específicos	14
1.6. Propuesta.....	15
1.7. Hipótesis	16
1.7.1. <i>Hipótesis general</i>	16
1.7.2. <i>Hipótesis específicas</i>	16
Tabla 1: Matriz de consistencia	18
1.8. Organización de la investigación	19
1.9. Producción científica.....	20
CAPÍTULO 2. MARCO TEÓRICO	21
2.1. Marco filosófico o epistemológico.....	21
2.2. Antecedentes del problema	22
2.3. Bases Teóricas.....	24

2.3.1. Citología.....	24
2.3.2. Examen citológico.....	26
2.3.3. Derrame <i>Pleural</i>	27
2.3.4. <i>Aprendizaje profundo</i>	31
2.3.5. <i>Redes</i> neuronales Convolucionales	33
2.3.6. <i>Arquitectura</i> ResNet50V2	38
2.3.7. <i>Arquitectura</i> DesNet121	42
2.3.8. <i>Arquitectura</i> Inception.....	43
2.3.9. Aprendizaje <i>por</i> transferencia	44
2.3.10. Clasificador de votación	46
Situación del recurso humano MEDICO ESPECIALISTA en el MINSA	47
CAPÍTULO 3. ESTADO DEL ARTE	48
3.1. Selección de artículos de trabajos relacionados	48
3.2. Desarrollo.....	50
3.3. Resultados de la revisión	58
3.4. Comentarios de la revisión.....	65
CAPÍTULO 4. METODOLOGÍA DE LA INVESTIGACIÓN	67
4.1. Metodología general de la investigación	68
Tamaño de muestra.....	69

Selección de muestra.....	69
Técnicas de recolección de datos.....	69
Análisis e interpretación de la información.....	70
4.2. Diseño del Artefacto: Ciencia del Diseño.....	70
4.2.1. Identificación del problema y sus motivaciones	73
4.3. Justificación de la metodología DSR	75
CAPÍTULO 5. CONSTRUCCIÓN DEL CONJUNTO DE DATOS DE	
IMÁGENES CITOLÓGICAS DE LÍQUIDO PLEURAL	77
5.1. Construcción del conjunto de datos de imágenes citológicas de líquido pleural.	78
5.2. Captura de imágenes.....	80
5.3. Categoría diagnóstica de las imágenes	80
5.4. Preprocesamiento de las imágenes	81
5.4.1. Etiquetado de imágenes para organizar el conjunto de datos	81
5.4.2. Dimensionamiento de las imágenes	83
5.4.3. Conversión a blanco y gris.....	83
5.5. Cumplimiento de normativas éticas y legales.....	87
CAPÍTULO 6. DISEÑO Y DESARROLLO DEL ARTEFACTO	88
6.1. Desarrollo de experimentos	90
6.2. Generación de nuevas imágenes	90

6.3.	Creación de escenarios	94
6.4.	Experimentos	94
6.5.	Modelado para clasificar imágenes	96
6.6.	Distribución del conjunto de datos.....	99
6.7.	Entrenamiento y validación	100
CAPÍTULO 7. RESULTADOS Y DISCUSIÓN.....		111
7.1.	Evaluación por escenarios	111
7.2.	Resultados	114
7.3.	Discusión de resultados	119
CAPÍTULO 8. CONCLUSIONES Y TRABAJOS FUTUROS.....		122
8.1.	Conclusiones	122
8.2.	Limitaciones.....	124
8.3.	Trabajos futuros.....	125
8.4.	Recomendaciones.....	125
CONTRIBUCIONES		127
ANEXOS		144

LISTA DE FIGURAS

Figura 1. Derrame Pleural	2
Figura 2. Tendencia de casos registrados de cáncer, Perú, 2019-2023	6
Figura 3. Aplicaciones de la citología	25
Figura 4. Pulmón normal y pulmón con Derrame Pleural	28
Figura 5. Sospecha de DPM	29
Figura 6. Métodos de diagnóstico	31
Figura 7. Cuatro topologías de red neuronal	33
Figura 8. Estructura de una CNN	34
Figura 9. Capa de convolución (filtro, entrada) y mapa de características	35
Figura 10. Matriz de entrada con max-pooling y average-pooling	36
Figura 11. Imágenes al azar de Imagenet	38
Figura 12. Bloque residual	41
Figura 13. Redes más profundas	41
Figura 14. Arquitectura DesNet121	42
Figura 15. Arquitectura InceptionV3	43
Figura 16. Transfer Learning	45
Figura 17. Voting classifier	46
Figura 18. Informe RRHH del sector salud al 2023 publicado en 2024	47

Figura 19. Selección de artículos	51
Figura 20. Estrategia de búsqueda de información	51
Figura 21. Taxonomía de la revisión de la literatura.....	57
Figura 22. Metodología de Ciencia del Diseño DSR	73
Figura 23. Microscopio óptico del HNCH.....	79
Figura 24. Código para organizar archivos	82
Figura 25. Imágenes renombradas y numeradas.....	82
Figura 26. Imágenes del conjunto de datos en escala de grises	84
Figura 27. Imágenes nombradas y numeradas por categoría diagnóstica.....	85
Figura 28. <i>Imágenes del conjunto de datos construido.</i>	86
Figura 29. Modelado propuesto.....	89
Figura 30. Parámetros del ImageDataGenerator para la transformación de imágenes.	92
Figura 31. Efectos aplicados en: A vertical_flip=True, B horizontal_flip=True, C rotation_range=30 y D width_shift_range=0.15.	92
Figura 32. <i>Imágenes del conjunto de datos Body cavity fluid cytology images: a) NFM- Negativa para malignidad y b) MAL-Malignas.</i>	96
Figura 33. Arquitectura de red neuronal convolucional utilizando a) RedNet50V2, b)DesNet131 y c) InceptionV3	98
Figura 34. Resultado del entrenamiento y validación para ResNet50V2 escenario1	102

Figura 35. Resultado del entrenamiento y validación para Desnet121 escenario1 .	102
Figura 36. Resultado del entrenamiento y validación para Inception escenario1 ..	103
Figura 37. Resultado del entrenamiento y validación para ResNet50V2 escenario 2.	105
Figura 38. Resultado del entrenamiento y validación para DesNet121 escenario 2	106
Figura 39. Resultado del entrenamiento y validación para Inception escenario 2 .	107
Figura 40. Resultado del entrenamiento y validación para ResNet50V2 escenario 3.	108
Figura 41. Resultado del entrenamiento y validación para DesNet121 escenario 3	109
Figura 42. Resultado del entrenamiento y validación para Inception escenario 3 .	110
Figura 43. Evaluación escenario 1: (A) ResNet50V2, (B) DesNet21 y (C) InceptionV3	111
Figura 44. Evaluación escenario 2: (A) ResNet50V2, (B) DesNet21 y (C) InceptionV3	112
Figura 45. Evaluación del escenario 3: (A) ResNet50V2, (B) DesNet12 y (C) InceptionV3	113

LISTA DE TABLAS

Tabla 1. Aplicaciones de la citología en la medicina	26
Tabla 2. Estrategia PICO	49
Tabla 3. Criterio de búsqueda	50
Tabla 4. Análisis de los hallazgos	58
Tabla 5. Evaluación de los artículos seleccionados	63
Tabla 6. Categorías diagnósticas de acuerdo con el TIS.	81
Tabla 7. Imágenes nombradas y numeradas por categoría diagnóstica	86
Tabla 8. Estructura del conjunto de datos	87
Tabla 9. Categorías diagnosticas para la investigación.	90
Tabla 10. Categoría diagnóstica y número de imágenes del conjunto de datos HNCH	95
Tabla 11. Categoría diagnóstica y número de imágenes del conjunto de datos KAGGLE	95
Tabla 12. Número de imágenes al inicio del experimento	99
<i>Tabla 13. Número de imágenes generadas al 50%</i>	99
Tabla 14. Número de imágenes generadas al 300%	100
Tabla 15. Parámetros y resultados obtenidos en el entrenamiento y validación del escenario 1.	101

Tabla 16. Parámetros y resultados obtenidos en el entrenamiento y validación del escenario 2	104
Tabla 17. Parámetros y resultados obtenidos en el entrenamiento y validación del escenario 3	107
Tabla 18. Métrica de evaluación de los modelos del escenario 1 en el conjunto de datos propio y en el conjunto de datos de Kaggle.....	115
Tabla 19. Métrica de evaluación de los modelos del escenario 2 en el conjunto de datos propio y en el conjunto de datos de Kaggle.....	117
Tabla 20. Métrica de evaluación de los modelos del escenario 3 en el conjunto de datos propio y en el conjunto de datos de Kaggle.....	118

RESUMEN

El Derrame Pleural, es una patología que se manifiesta con la acumulación excesiva o fuera del rango normal de líquido pleural en la zona que rodea los pulmones. La citología es una rama de la medicina que estudia las células en sus diferentes aspectos morfológicos, funcionales y genéticos; desde microscopios simples hasta técnicas modernas, la evolución continúa, permitiendo su uso en oncología y análisis celular.

El examen citológico es una técnica diagnóstica muy utilizada y de alta precisión para confirmar enfermedades y determinar su estadio. El proceso consiste en identificar signos atípicos, como células con formas y texturas irregulares. El diagnóstico médico se complica debido a la variabilidad de los estados patológicos en las muestras de líquido pleural, que pueden incluir enfermedades inflamatorias, infecciosas y neoplasias. La detección temprana de malignidad es crucial para un tratamiento oportuno, reduciendo el sufrimiento y potencialmente prolongando la vida del paciente. Además, la escasez de médicos patólogos y la demora en los diagnósticos abren oportunidades para la implementación de técnicas de inteligencia artificial, las cuales han demostrado un impacto significativo en la medicina diagnóstica en los últimos años. La investigación se llevó a cabo utilizando la metodología de Investigación en Ciencias del Diseño (DSR), lo que permitió un diseño estructurado, una implementación efectiva y una validación exhaustiva de la técnica propuesta.

Se construyó un conjunto de datos de imágenes citológicas de líquido pleural que contiene 2640 imágenes de muestras de citología de líquido pleural. Para identificar células neoplásicas, se diseñó un modelo basado en Redes Neuronales Convolucionales (CNN), utilizando Transfer Learning y tres arquitecturas de CNN pre entrenadas ResNet50V2, DesNet121 e InceptionV3. Se evaluó el desempeño al identificar células MAL (malignas) y células NFM (negativas para malignidad - benignas) según la clasificación del sistema internacional (TIS) para informar citología de líquidos serosos: riesgo de malignidad (MAL y NFM).

Los tres modelos fueron entrenados, validados y evaluados con el mismo conjunto de datos, se evaluó el desempeño por modelo y se realizaron comparaciones combinadas mediante el Voting Classifier, para determinar qué modelo presenta la mejor precisión en la clasificación de imágenes citológicas.

Los algoritmos de aprendizaje profundo propuestos para clasificar imágenes citológicas de líquido pleural lograron un rendimiento superior al 90% en pruebas. Las combinaciones más efectivas fueron ResNet+Inception y ResNet+DesNet, la primera combinación obtuvo mayor precisión (0.94) pero fue menos generalizable con un conjunto de datos externo, a diferencia la segunda combinación, aunque obtuvo una menor precisión (0.92) fue más generalizable cuando se probó en un conjunto de datos externo. Los resultados confirman el potencial del método propuesto para ayudar a los patólogos a lograr un diagnóstico más preciso.

Palabras clave: Citología, Examen citológico, Derrame Pleural, Aprendizaje profundo, Red Neuronal Convolutacional (CNN), Aprendizaje por transferencia, Clasificador de votación.

ABSTRACT

Pleural effusion is a pathology that manifests as the excessive or abnormal accumulation of pleural fluid in the area surrounding the lungs. Cytology is a branch of medicine that studies cells in their different morphological, functional, and genetic aspects; from simple microscopes to modern techniques, evolution continues, allowing its use in oncology and cellular analysis.

Cytological examination is a widely used and highly accurate diagnostic technique for confirming diseases and determining their stage. The process involves identifying atypical signs, such as cells with irregular shapes and textures. Medical diagnosis is complicated by the variability of pathological states in pleural fluid samples, which can include inflammatory, infectious, and neoplastic diseases. Early detection of malignancy is crucial for timely treatment, reducing suffering and potentially prolonging the patient's life. Furthermore, the shortage of pathologists and the delay in diagnoses open opportunities for the implementation of artificial intelligence techniques, which have demonstrated a significant impact on diagnostic medicine in recent years. The research was conducted using the Design Science Research (DSR) methodology, which allowed for a structured design, effective implementation, and thorough validation of the proposed technique.

A dataset of pleural fluid cytology images was constructed, containing 2,640 images of pleural fluid cytology samples. To identify neoplastic cells, a Convolutional Neural Network (CNN)-based model was designed using Transfer Learning and three pre-trained CNN architectures: ResNet50V2, DesNet121, and InceptionV3. Performance was evaluated in identifying MAL (malignant) cells and NFM (negative for malignancy - benign) cells according to the International System for Reporting Serous Fluid Cytology (TIS) classification: risk of malignancy (MAL and NFM). The three models were trained, validated, and evaluated on the same dataset. Model performance was assessed, and combined comparisons were made using the Voting Classifier to determine which model exhibits the best accuracy in classifying cytological images. The proposed deep learning

algorithms for classifying pleural fluid cytological images achieved performance above 90% in testing. The most effective combinations were ResNet+Inception and ResNet+DesNet. The first combination achieved higher accuracy (0.94) but was less generalizable to an external dataset. The second combination, while lower in accuracy (0.92), was more generalizable when tested on an external dataset. The results confirm the potential of the proposed method to help pathologists achieve more accurate diagnoses.

Keywords: Cytology, Cytological Examination, Pleural Effusion, Deep Learning, Convolutional Neural Network (CNN), Transfer Learning, Voting Classifier.

CAPÍTULO 1. INTRODUCCIÓN

Los modelos tecnológicos basados en técnicas de Inteligencia Artificial (IA) están en tendencia creciente en el área de la medicina, sobre todo en actividades que requieren resultados precisos para abordar problemáticas complejas mediante el desarrollo e implementación de soluciones tecnológicas innovadoras, como en los diagnósticos de enfermedades y apoyo al médico con resultados precisos. Muchas veces el médico toma decisiones basadas en la experticia y conocimientos por lo que, debido a motivos de complejidad de información puede errar, comprometiendo la salud humana. Los diagnósticos que requieren de análisis sofisticados, considerando sus niveles de evolución, como la clasificación de células neoplásicas –que sin el tratamiento médico oportuno– puede evolucionar poniendo en riesgo la vida del paciente. Con este modelo se pretende que el médico pueda aún tratar eficazmente para salvar el paciente. La presente investigación se enmarca en la intersección entre los sistemas inteligentes, el procesamiento de imágenes médicas y la necesidad crítica de optimizar los recursos sanitarios en el Perú, demostrando cómo la tecnología puede servir como catalizador para resolver desafíos estructurales del sector salud.

La pleura es una membrana serosa que cubre los pulmones. Cuando en la Pleura se produce una acumulación de líquido excesiva o por encima del límite normal, significa que se ha producido una condición patológica conocida como Derrame Pleural (DP), como se observa en la Figura 1. El DP es una de las patologías más frecuentes, que incluye enfermedades sistémicas, mal funcionamiento de órganos y el uso de medicamentos (Light, 2002).

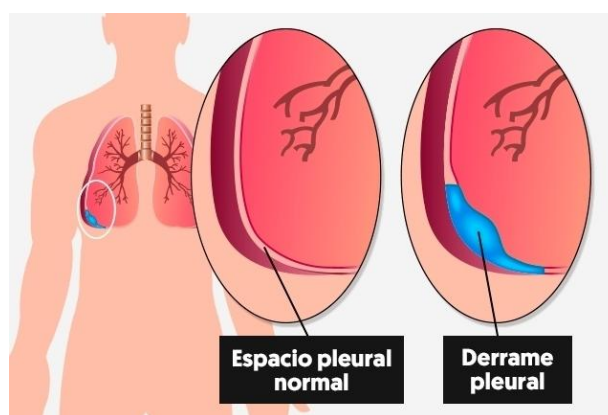


Figura 1. Derrame Pleural

Fuente. Tomada de <https://neumologoencdmx.com/enfermedades/derrame-pleural/>

Existen muchas causas que origine un DP, que pueden estar o no relacionadas con alguna enfermedad de los pulmones, algunas infecciones, cáncer, enfermedades del corazón, hígado, riñones, etc. Si el DP está comprometido con neoplasia, se clasifica en: Derrame Pleural Benigno (DPB) y Derrame Pleural Maligno (DPM) (Bielsa et al., 2008). Para determinar la malignidad se realiza un examen citológico de líquido pleural. El examen citológico se realiza para identificar células cancerosas y precancerosas, así como también para identificar otras afecciones (Rivera, 2022).

El Derrame Pleural es una de las causas de muertes en el mundo y representa el 50% de pacientes con diagnóstico de cáncer. La detección temprana y un diagnóstico preciso es fundamental para mejorar la posibilidad de una atención a tiempo y calidad de vida para los pacientes.

La citología es una rama de la medicina que a través del examen citológico permite analizar el líquido pleural y detectar posibles anomalías que indicarían señales de cáncer. Sin embargo, interpretar la imagen no es una labor sencilla, implica una revisión exhaustiva, repetición y subjetividad del especialista patólogo, lo que puede facilitar el riesgo a un error humano. Por esta razón, se hace necesario encontrar soluciones que faciliten y mejoren el diagnostico de los exámenes de líquido pleural.

El uso de técnica de Inteligencia Artificial (IA) en la detección y diagnóstico de cáncer en los últimos años ha sido tema de investigación. Los estudios realizados han explorado y demostrado la eficacia de diversas arquitecturas de modelos de aprendizaje profundo para analizar imágenes citológicas, con el único objetivo de apoyar el diagnóstico médico, mejorar la precisión y eficiencia en la detección de células cancerosas en líquido pleural.

Los trabajos analizados evidencian el interés creciente en aplicar redes neuronales profundas para la detección de células malignas en imágenes citológicas de líquido pleural. Sin embargo, persisten limitaciones frecuentes: a) cohortes pequeñas o desbalanceadas (Sanyal & Dey, 2023), b) dependencia de etiquetas débiles a nivel portaobjetos (Xie et al., 2022), y c) no realizan validación externa. Además, si bien se han logrado métricas destacables ($AUC > 0.90$ en varios casos), reconocen la necesidad de modelos más interpretables y generalizables, así como de incorporar aumento de datos para enfrentar la escasez de imágenes.

En contraste, el modelo propuesto en este trabajo se centra en transfer learning con ResNet50V2, DenseNet121 e Inception, evaluado bajo diferentes escenarios de aumento de datos (50 % y 300 %), lo que permite analizar cómo esta técnica impacta en la robustez del modelo. Este enfoque complementa y extiende los estudios previos, aportando evidencia sobre la efectividad del aumento de datos para mejorar la clasificación de células benignas (NFM) y malignas (MAL) en contextos de datos limitados.

1.1. Antecedentes

El Derrame Pleural, definido como la acumulación anormal de líquido en el espacio pleural, ha sido objeto de estudio desde la antigüedad debido al impacto con la respiración. En la medicina hipocrática, se identificaron los primeros signos clínicos asociados al DP, como el ruido torácico y la disminución de los sonidos respiratorios, lo que marcó el inicio del diagnóstico basado en la observación y la palpación (Light, 2021). Con el desarrollo de la medicina moderna en el siglo XIX, las técnicas como la percusión de Auenbrugger

y la auscultación de Laënnec permitieron avanzar en la evaluación clínica de esta condición (Hooper et al., 2010).

El diagnóstico del DP evolucionó en el siglo XX con la introducción de las radiografías de tórax, que proporcionaron una imagen clara de los cambios pleurales. Posteriormente, herramientas como la ecografía torácica y la tomografía computarizada permitieron una visualización más detallada del líquido pleural y de la posible etiología subyacente (Porcel, 2020). La toracocentesis, técnica introducida en el mismo período, se consolidó como un método diagnóstico clave, permitiendo diferenciar entre derrames transudativos y exudativos mediante criterios bioquímicos como los de (Light, 2002).

En la actualidad, el estudio del DP se centra en integrar métodos tradicionales con técnicas avanzadas, como el análisis citológico, para mejorar la precisión diagnóstica y personalizar los tratamientos según la condición subyacente del paciente (Porcel, 2020). Se resalta la importancia de continuar investigando en esta área para abordar de manera integral uno de los desafíos clínicos más comunes en la práctica médica.

A nivel mundial, el derrame pleural afecta aproximadamente a 1.5 millones de personas al año solo en los Estados Unidos, y su incidencia es aún mayor en regiones con alta prevalencia de infecciones como la tuberculosis (Hooper et al., 2010). Además, existe una relación destacada entre el DP y el cáncer, ya que hasta el 40% de los pacientes con neoplasias avanzadas desarrollan esta condición, siendo el cáncer de pulmón y el de mama las causas malignas más comunes (Porcel, 2018).

En países en desarrollo, donde no hay investigaciones y las instalaciones de salud son inadecuadas y el costo del tratamiento es alto, el examen citológico de los fluidos corporales se considera una herramienta económica, rápida y altamente efectiva para identificar la etiología del derrame con la predicción de la enfermedad subyacente. El examen citológico puede detectar células malignas en los derrames, muchas veces es superior a la biopsia para detectar la malignidad. Una historia médica bien realizada, antecedentes de exposición al asbesto, ausencia de otras neoplasias malignas en otras partes del cuerpo y hallazgos radiológicos son ayudas muy importantes en la confirmación

del diagnóstico citológico de malignidad. A veces, el diagnóstico definitivo no se alcanza solo con la citología y la necesidad de una biopsia de tejido o incluso otros métodos auxiliares se vuelve obligatoria (El-Sheikh, 2012).

De acuerdo con la Organización Mundial de la Salud OMS (2020), el cáncer se registra como la principal causa de muerte en el mundo y en el año 2020 aproximadamente se registraron 10 millones de fallecidos, lo que equivale a una de cada seis personas con la enfermedad. Los tipos de cáncer más comunes son los de mama, pulmón, colon y recto y próstata. En Estados Unidos en el año 2020, se diagnosticó 1 806 590 casos nuevos de cáncer y que 606 520 personas morirían por la enfermedad (Instituto Nacional del Cáncer, 2020).

La conexión entre el cáncer y el derrame pleural es significativa, ya que los derrames pleurales malignos DPM son consecuencias de una complicación común en pacientes con cáncer avanzado. Se ha estimado que entre el 15% y el 50% de los pacientes con cáncer desarrollan DPM en el transcurso de la enfermedad (Shojaee et al., 2021)

Los avances en el tratamiento han reducido la mortalidad por neoplasias malignas. Sin embargo, el progreso se desacelera en cánceres detectables tempranamente mediante pruebas de cribado. Preocupan las disparidades raciales, socioeconómicas y geográficas en cánceres prevenibles, como el de cuello uterino y pulmón. Es fundamental aumentar la inversión para garantizar una aplicación equitativa y generalizada de las intervenciones actuales, así como fortalecer la investigación básica y clínica para mejorar el conocimiento y desarrollar nuevos tratamientos (Siegel et al., 2021).

De acuerdo con el Centro Internacional de Investigaciones sobre el Cáncer -CIIC, (2024) se estima más de 35 millones de casos nuevos de neoplasias en 2050, significa que aumentará en 77% con la estimación de 20 millones de casos del 2022. Este crecimiento de casos nuevos de cáncer a nivel mundial se debe en parte al envejecimiento y al crecimiento de la población, también a la exposición a factores de riesgo asociados al desarrollo socioeconómico. Se incluye el tabaco, el alcohol y la obesidad como factores claves del crecimiento de casos nuevos de cáncer y la contaminación ambiental como un

factor importante de riesgo medioambientales. Si relacionamos estas cifras con las estadísticas de los DPM, 50% de los DPM corresponden a neoplasias, se infiere que los centros de salud específicamente el área de Citología tendrá mayor carga de muestras para diagnósticos de DPM debido a las metástasis.

En el Perú, de acuerdo a (MINSA, 2023a) el número de casos de cáncer según el Registro de Cáncer Activo aumentó en la DIRIS Lima Centro, DIRIS Lima Este y DIRIS Lima Norte, lo que puede traducirse en una alta tasa de morbilidad y mortalidad. Según el boletín (MINSA, 2023b) en el año 2023, se registraron 13899 casos de cáncer; de los cuales 4970 correspondieron a hombres y 8929 a mujeres, como se muestra en la Figura 2.

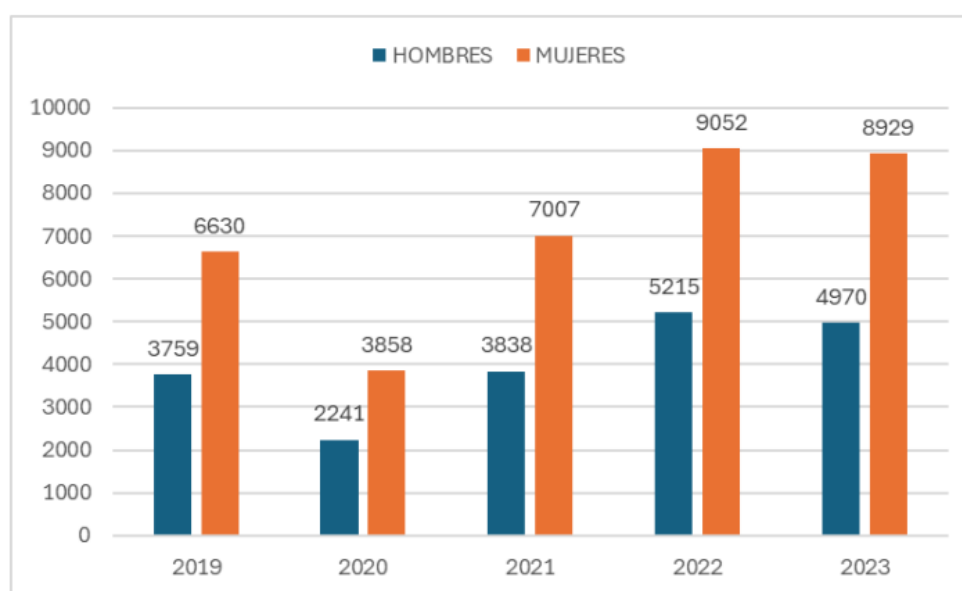


Figura 2. Tendencia de casos registrados de cáncer, Perú, 2019-2023

Según Ministerio de Salud & Centro Nacional de Epidemiología (2019), para lograr la condición básica del 3er objetivo de desarrollo sostenible (ODS), Salud y Bienestar y la Cobertura Universal en Salud, se debe contar con una cantidad suficiente de trabajadores

de salud, capacitados y distribuidos para poder asegurar el acceso equitativo a los servicios de salud.

En el 2021, hubo 14072 médicos especialistas, concentrados en Lima 6715 y los tres departamentos que tienen menos de 100 médicos especialistas son Madre de Dios, Pasco y Huancavelica con 37, 58 y 89 respectivamente (Centro Nacional de Epidemiología, 2021).

En el Perú, existen hospitales y centros de salud, que carecen de médicos especialistas y apenas cuentan con personal técnico sintiéndose la ausencia de un profesional de la salud. Ante la carencia de médicos especialistas, en particular de patólogos para diagnosticar y clasificar el Derrame Pleural, se puede implementar técnicas de IA que sirvan de apoyo médico. Se observaron las desigualdades en salud de la población que habita en zonas rurales de la sierra, de la selva y en aquellos pueblos que tienen como lengua nativa quechua, aymara o lenguas amazónicas, reciben el mayor impacto de la pobreza (Oyola, n.d.). Las principales barreras para implementar planes o estrategias nacionales de salud es la capacidad insuficiente de recursos humanos y tecnologías, herramientas y métodos insuficientes (Who, 2023).

1.2. Situación problemática

La citología es una técnica muy usada y eficaz para el diagnóstico de DPM, se distingue por su facilidad, rentabilidad y rapidez de los resultados (Dermawan & Policarpio-Nicolas, 2020). Los resultados a tiempo de la prueba citológica de derrame pleural contribuyen con una atención rápida, logrando acceso a tratamiento y mejora en general del paciente. Debido a las limitaciones del estudio citológico que incluye la parte subjetiva de especialista y un posible error humano en el diagnóstico, en los últimos años, investigadores vienen explorando enfoques diferentes y alternativos que puedan dar resultados más precisos y fiables. Las imágenes de diapositivas, láminas o portaobjetos, viene hacer el insumo principal en el cual podemos aplicar la Inteligencia Artificial (IA) en la Patología Digital (Carmona, 2021) .

En el Perú, el acceso a los servicios de salud en condiciones iguales es un problema persistente, sobre todo en el diagnóstico y tratamiento de enfermedades complejas como el DP asociado a células neoplásicas. La falta de médicos especialistas como patólogos en provincias es crítica. Según el (CMP, 2021), más del 60% de los especialistas se concentran en Lima, mientras que las regiones más alejadas carecen de cobertura adecuada. Esta desigualdad geográfica limita el acceso oportuno a procedimientos esenciales como el examen citológico y la biopsia, retardando el diagnóstico y el inicio del tratamiento en pacientes con sospecha de enfermedades malignas (MINSA del Perú, 2021). Además, las estadísticas estiman que entre el 30% y el 50% de los casos de derrame pleural tienen origen maligno, según datos del (INEN, 2022). La falta de diagnósticos tempranos en estos pacientes reduce significativamente sus tasas de supervivencia (Light, 2021).

Estas necesidades nos abren la posibilidad de encontrar soluciones que puedan atender esta falencia, y revisando la literatura encontramos que desde hace varios años se viene desarrollando trabajos que han incorporado la IA a la medicina siendo una buena y precisa ayuda para diagnosticar determinadas enfermedades al igual que los médicos, la diferencia es que lo hacen en menos tiempo. Investigadores desarrollan y utilizan modelos de inteligencia artificial para detectar cáncer como por ejemplo el cáncer de mama, identificar patologías de la retina, o diagnosticar neumonía. Estos modelos además de realizar el diagnóstico correcto proporcionan valoraciones de la enfermedad y hasta recomiendan un tratamiento de acuerdo con la patología diagnosticada.

Con el desarrollo acelerado de la tecnología, nos encontramos en un escenario positivo para la medicina, especialmente en el área de diagnóstico por imágenes. Esto nos lleva a investigar con la IA mediante técnicas de Aprendizaje Profundo para diagnosticar imágenes citológicas y estudiar aquellas debilidades e inconvenientes de la citología tradicional.

Por lo tanto, la implementación de un diagnóstico automatizado de células neoplásicas en líquido pleural representa una solución innovadora y necesaria. Este enfoque no solo mitigaría el déficit de especialistas, sino que también contribuiría a reducir las inequidades

en salud, facilitando diagnósticos más rápidos y precisos, incluso en regiones con recursos limitados.

Diagnosticar a tiempo células neoplásicas en imágenes citológicas del líquido pleural es un reto para la medicina. Con el desarrollo de la inteligencia artificial, los modelos de aprendizaje profundo han demostrado un rendimiento prometedor en la clasificación de imágenes médicas. En este contexto, las redes neuronales convolucionales (CNN) y el aprendizaje por transferencia han surgido como herramientas clave para mejorar la precisión del diagnóstico automatizado.

Esta investigación, hasta donde se ha tomado conocimiento, es el primer estudio con un conjunto de datos propios en el Perú. Se encontraron estudios internacionales, como en (Win et al., 2018) que experimentaron con conjuntos de datos muy pequeños de 125 imágenes citológicas. (B. Baykal et al., 2020) experimentaron con 120 imágenes citológicas, ambas investigaciones sugieren el aumento de datos para cubrir variabilidad clínica y probar arquitecturas más recientes. En el Perú, debido a que la parte asistencial del sistema de salud nacional no está digitalizado se presenta la oportunidad para implementar la automatización del examen citológico.

Los estudios previos revisados presentan limitaciones de conjuntos de datos pequeños y desbalanceados, validación interna sin generalización externa, ausencia de ensembles y falta de análisis sistemático del fine-tuning. Estas brechas constituyen la problemática técnica que aborda la presente tesis doctoral.

1.2.1. Importancia

En el Perú, el problema de salud pública, cuyo objetivo es brindar atención médica oportuna, es de alta incidencia y este proyecto cubrirá la brecha de especialistas Ministerio de Salud del Perú (2021).

1.2.2. Novedad

La tecnología aplicada en las diferentes especialidades de la medicina puede considerarse una novedad en especial para el diagnóstico de células cancerosas. La complejidad requiere un enfoque innovador y la incorporación de tecnologías emergentes y el procesamiento de imágenes. Además, la investigación colaborativa y multidisciplinaria entre expertos médicos, informáticos y otras disciplinas, donde aportan sus conocimientos y habilidades para desarrollar una solución integral.

1.2.3. Interés

Los investigadores ven en esta área un campo fértil para la innovación y la búsqueda de soluciones efectivas para esta patología que resulta atractivo para los profesionales y empresas tecnológicas que buscan contribuir a mejorar de la atención médica, el bienestar de las personas y las tasas de supervivencia de pacientes.

También se ocasiona un impacto profundo en la salud pública ya que las entidades de salud y los responsables de las políticas públicas muestran interés en reducir la carga de trabajo en el sistema de salud. La aplicación de un modelo de IA podría tener un impacto positivo en la salud pública al acelerar el tiempo del diagnóstico y cubrir la brecha de la falta de médicos especialistas, así como superar las barreras geográficas y económicas en el acceso a la atención médica independientemente de su ubicación geográfica o situación socioeconómica.

1.2.4. Viabilidad

Debido a los avances tecnológicos y disponibilidad de datos hacen que el desarrollo de modelos de IA sea cada vez más viable. Las instituciones de salud y centros de investigación pueden recopilar y almacenar datos para entrenar, probar y validar el modelo de IA. La colaboración entre instituciones de salud, universidades y empresas tecnológicas puede aumentar la viabilidad del proyecto. La combinación de recursos, experiencia y conocimientos de diferentes entidades puede acelerar el desarrollo y la implementación del modelo de IA para el diagnóstico de células cancerosas en líquido pleural.

1.3. Formulación del Problema

La complejidad de los indicios de células neoplásicas en los exámenes citológicos de líquido pleural generalmente puede ser interpretada por los médicos como células malignas o células benignas. A pesar de la evolución de sistemas computacionales inteligentes existen hospitales que no cuentan con un modelo computacional que sirva de apoyo en el diagnóstico.

1.3.1. Problema principal

Para el desarrollo de este trabajo se plantea la pregunta de investigación:

¿Cómo el desarrollo de un modelo basado en aprendizaje profundo sirve de apoyo médico y brinda un resultado preciso, en la clasificación de células neoplásicas en imágenes de citología de líquido pleural?

1.3.2. Problemas secundarios

PS1: ¿Qué estudios de investigación se han realizado usando técnicas IA para la clasificación de células neoplásicas en imágenes de citología de líquido pleural?

PS2: ¿Cómo construir un conjunto de datos locales de imágenes citológicas de líquido pleural?

PS3: ¿Cómo implementar un modelo inteligente que permita clasificar células neoplásicas en imágenes citológicas de líquido pleural?

PS4: ¿Cómo generalizar el modelo verificando la precisión en un conjunto de datos nuevo?

1.4. Justificación

Este proyecto se encuentra dentro de la línea de investigación de la Inteligencia Artificial del Doctorado de Ingeniería de Sistemas e Informática de la Universidad Nacional Mayor de San Marcos, que se describe en el plan de estudios.

El impacto de esta investigación radica en la posibilidad de mejorar la precisión y rapidez del diagnóstico de células neoplásicas en imágenes citológicas de líquido pleural, otorgando una herramienta automatizada que pueda asistir a los médicos en la detección de células neoplásicas. Así mismo, la construcción de un conjunto de datos exclusivo para esta investigación representa un avance significativo en la disponibilidad de datos para futuras investigaciones en IA aplicada a la citología.

Por lo tanto, la presente investigación se justifica porque:

- a) El sector salud en el Perú no posee la cantidad suficiente de especialistas para atender a la población, especialmente en diagnósticos de gran nivel de complejidad y muy especializado para diagnosticar células neoplásicas en citología de líquido pleural.
- b) El estudio propone el desarrollo de un modelo inteligente, que integra metodologías de IA con interfaces del microscopio, la cámara incorporada del microscopio y un conjunto de datos propio de imágenes, para clasificar células neoplásicas a partir de imágenes citológicas de líquido pleural, lo cual constituye un aporte tecnológico a las soluciones existentes.
- c) La construcción de un conjunto de datos locales y el trabajo con los expertos permitirá validar el algoritmo que diagnostica las células cancerosas en las imágenes citológicas.
- d) La implementación del algoritmo para diagnosticar células neoplásicas permitirá que el paciente sea derivado inmediatamente para confirmación médica del diagnóstico y prescribirle un tratamiento, por lo que mejorará la calidad del servicio a la población que pertenece a los establecimientos de salud nivel.
- e) También se pretende que este proyecto sea un punto de partida para trabajos futuros e implementar nuevas herramientas para estudiar, detectar y diagnosticar otras patologías mediante imágenes usando técnicas de aprendizaje profundo.

Según los autores (Bravo et al., 2021a), consideran que el uso de redes convolucionales pueden ser usadas como complemento para mejorar la eficiencia en las herramientas del diagnóstico asistido del cáncer.

1.4.1. Justificación Teórica

Esta investigación se realiza con el propósito de aportar al conocimiento existente sobre el uso de la tecnología en el diagnóstico de células neoplásicas en líquido pleural (Derrame Pleural Maligno) y su implicancia en las patologías del paciente, con un diagnóstico eficaz, preciso y una validación por expertos. Los resultados de esta investigación pudieran tomarse en cuenta como aporte al conocimiento del área de informática, ya que se estaría demostrando que a través del algoritmo de diagnóstico de imágenes citológicas permitirá mejorar el diagnóstico y otorgar el tratamiento que corresponda a los pacientes.

1.4.2. Justificación Práctica

Desde el punto de vista práctico se propone el diseño y aplicación de un modelo algorítmico, el cual permita el diagnóstico de células cancerosas que hubiere en el líquido pleural, así mismo para que cubra la brecha de especialistas anatomopatólogos en los establecimientos de salud.

1.4.3. Motivación del estudio

El diagnóstico citológico de líquido pleural para detectar células neoplásicas representa un desafío crítico en la práctica clínica actual. La interpretación de los resultados requiere alta especialización y presenta variabilidad entre los especialistas, generando inconsistencias diagnósticas que pueden retrasar el tratamiento oportuno de pacientes oncológicos. Además, la creciente demanda de estudios citológicos y la limitada disponibilidad de profesionales de la salud especializados originan retrasos que impactan en la atención médica. Los errores diagnósticos en este contexto pueden tener consecuencias graves, desde falsos negativos que retrasan el tratamiento hasta falsos positivos que generan procedimientos innecesarios.

El desarrollo de un modelo de aprendizaje profundo para clasificar células neoplásicas representa una oportunidad única para abordar estos desafíos. Las redes neuronales convolucionales han demostrado capacidades extraordinarias para identificar patrones morfológicos complejos en imágenes médicas, ofreciendo consistencia, reproducibilidad y escalabilidad.

Esta investigación está motivada por la posibilidad de mejorar la precisión diagnóstica, optimizar el flujo de trabajo en las instituciones de salud, y proporcionar una herramienta de apoyo que beneficie directamente a los pacientes mediante diagnósticos más rápidos y precisos, contribuyendo significativamente a la medicina de precisión y la salud pública.

1.5. Objetivos de la Investigación

1.5.1. Objetivo General

Desarrollar un modelo basado en Aprendizaje Profundo, usando la técnica de aprendizaje por transferencia y las arquitecturas ResNet50V2, DenseNet121 e InceptionV3, para clasificar de manera eficiente imágenes de células neoplásicas en citología de líquido pleural.

1.5.2. Objetivos Específicos

OE1: Analizar las diferentes técnicas y enfoques de los estudios realizados en Aprendizaje Profundo para mejorar la precisión del modelo.

OE2: Construir un conjunto de datos locales de imágenes citológicas de líquido pleural para el entrenamiento, prueba y validación del modelo.

OE3: Implementar tres escenarios para determinar la configuración más adecuada para optimizar la detección de células malignas

OE4: Validar la generalización del modelo verificando la precisión en un conjunto de datos nuevo.

1.6. Propuesta

El modelo de aprendizaje profundo propuesto se fundamenta en un flujo metodológico estructurado que inicia con la construcción del conjunto de datos de citología de líquido pleural, tal como se muestra en la Figura 3. Está conformado por imágenes digitales representativas de células benignas y malignas. Estas imágenes atraviesan un proceso de preprocesamiento, orientado a garantizar la calidad y homogeneidad de la información, que incluye la normalización de dimensiones, la conversión a escala de grises o ajustes de color, y la generación de nuevas instancias sintéticas mediante técnicas de aumento de datos (data augmentation).

Posteriormente, se establecen tres escenarios experimentales diferenciados por el grado de aumento de datos aplicado: (i) sin aumento, (ii) aumento del 50 %, y (iii) aumento del 300 %. En cada escenario se implementan modelos basados en redes neuronales convolucionales (CNN) con la técnica de aprendizaje por transferencia (transfer learning), empleando como arquitecturas base ResNet50V2, DenseNet121 e InceptionV3, reconocidas por su robustez en la extracción de características de imágenes biomédicas.

El conjunto de datos se divide en dos particiones: 70 % destinado al entrenamiento y validación de los modelos, y 30 % reservado para la evaluación final del desempeño. Durante el entrenamiento se adaptan y personalizan las CNN seleccionadas, ajustando sus capas superiores para la clasificación binaria de células benignas (NFM) y malignas (MAL).

La evaluación de los modelos se realiza mediante métricas ampliamente utilizadas en el ámbito del aprendizaje profundo para aplicaciones médicas, tales como exactitud (accuracy), precisión, exhaustividad (recall) y la medida F1. Este proceso de contrastación de resultados permite comparar el rendimiento de cada arquitectura bajo los distintos escenarios de aumento de datos, con el objetivo de determinar la configuración más adecuada para optimizar la detección de células malignas.

Finalmente, el modelo integra los resultados obtenidos y clasifica las imágenes en dos categorías diagnósticas: NFM (células benignas) y MAL (células malignas), aportando una herramienta computacional de apoyo al diagnóstico clínico en el análisis citológico del líquido pleural.

1.7. Hipótesis

1.7.1. Hipótesis general

Un modelo basado en Aprendizaje Profundo aplicando arquitecturas de Transfer Learning, entrenado con datos aumentados controlados (0%, 50%, 300%), mejorará significativamente la precisión y mantendrá $AUC \geq 0.85$ en un conjunto de datos externo.

1.7.2. Hipótesis específicas

HE1: Un modelo entrenado con datos aumentados controlados en tres escenarios (0%, 50%, 300%), mejorará significativamente la precisión y mantendrá $AUC \geq 0.85$ en un conjunto de datos externo.

HE2: La construcción del conjunto de datos locales de imágenes citológicas permite el entrenamiento, prueba y validación del modelo.

HE3: La implementación de escenarios permite obtener la configuración más adecuada para optimizar la detección de células malignas.

HE4: La generalización del modelo se verifica mediante la precisión alcanzada por un conjunto de datos nuevo.

Problema general ¿Cómo el desarrollo de un modelo basado en aprendizaje profundo sirve de apoyo médico y brinda un resultado preciso, en la clasificación de células neoplásicas en imágenes de citología de líquido pleural?	Objetivo general Desarrollar un modelo basado en Aprendizaje Profundo, usando la técnica de aprendizaje por transferencia y las arquitecturas ResNet50V2, DenseNet121 e InceptionV3, para clasificar de manera eficiente imágenes de células neoplásicas en citología de líquido pleural	Hipótesis general Un modelo basado en Aprendizaje Profundo aplicando arquitecturas de Transfer Learning, entrenado con datos aumentados controlados (0%, 50%, 300%), mejorará significativamente la precisión y mantendrá $AUC \geq 0.85$ en un conjunto de datos externo	Se trabajó con la Metodología de Ciencia del Diseño - MDSR	Determinar si se cumplieron los objetivos del diseño.
Problema específico 1 ¿Qué estudios de investigación se han realizado usando técnicas IA para la clasificación de células neoplásicas en imágenes de citología de líquido pleural?	Objetivo específico 1 Analizar las diferentes técnicas y enfoques de los estudios realizados en Aprendizaje Profundo para mejorar la precisión del modelo.	Hipótesis específica 1 Un modelo entrenado con datos aumentados controlados en tres escenarios (0%, 50%, 300%), mejorará significativamente la precisión y mantendrá $AUC \geq 0.85$ en un conjunto de datos externos.	Revisión sistemática para encontrar técnicas de IA para clasificar imágenes médicas. Contrastar técnicas y revisar sus resultados.	Identificar las características citológicas presentes en el líquido pleural MAL y NFM
Problema específico 2	Objetivo específico 2 Construir un conjunto de datos locales de imágenes citológicas de líquido	Hipótesis específica 2 La construcción del conjunto de datos locales de imágenes	Revisión sistemática para ubicar conjunto de datos	Revisión de la clasificación de las células por especialistas

¿Cómo construir un conjunto de datos locales de imágenes citológicas de líquido pleural?	pleural para el entrenamiento, prueba y validación del modelo.	citológicas permite el entrenamiento, prueba y validación del modelo.	de citología de líquido pleural. Estudiar la disponibilidad de datos.	
Problema específico 3 ¿Cómo implementar un modelo inteligente que permita clasificar células neoplásicas en imágenes citológicas de líquido pleural?	Objetivo específico 3 Implementar tres escenarios para determinar la configuración más adecuada para optimizar la detección de células malignas	Hipótesis específica 3 La implementación de escenarios permite obtener la configuración más adecuada para optimizar la detección de células malignas.	Seleccionar arquitecturas, adaptar modelos y realizar experimentos.	Evaluar las arquitecturas seleccionadas para la clasificación de células neoplásicas. Analizar resultados de los 3 escenarios.
Problema específico 4 ¿Cómo generalizar el modelo verificando la precisión en un conjunto de datos nuevo?	Objetivo específico 4 Validar la generalización del modelo verificando las combinaciones en un conjunto de datos nuevo.	Hipótesis específica 4 La generalización del modelo se verifica mediante la combinación con precisión alta.	Comparar resultados que generalicen el modelo.	Análisis de la precisión de los 3 escenarios: Comparación, combinaciones, generalizables.

Matriz de consistencia

1.8. Organización de la investigación

El trabajo de investigación se estructura en seis capítulos. Incluye la construcción de un conjunto de datos de imágenes citológicas de líquido pleural; y el diseño, desarrollo y prueba del modelo computacional para diagnóstico médico, teniendo en cuenta las seis actividades de la metodología DSR.

En el capítulo 1 se presenta y formula la situación problemática, el objetivo general, los objetivos específicos y la motivación para el desarrollo de la investigación. Se relaciona con las actividades 1 y 2 de la metodología DSR.

En el capítulo 2, expone los antecedentes y las bases teóricas de los conceptos vinculados al estudio que sustentan este trabajo de investigación: Citología, examen citológico, aprendizaje profundo, redes neuronales convolucionales, aprendizaje por transferencia y clasificador de votación.

En el capítulo 3, se ha desarrollado el estado del arte respondiendo a las preguntas de investigación formuladas.

En el capítulo 4, se ha detallado la metodología de investigación utilizada en el trabajo siguiendo la normativa de la UNMSM

En el capítulo 5, se presenta la construcción del conjunto de datos de imágenes locales, realizado por un equipo multidisciplinario formado por el Departamento de Anatomía Patológica del Hospital Nacional Cayetano Heredia (Ver ANEXO 1 y ANEXO 2) y la Unidad de Posgrado de la Facultad de Ingeniería de Sistemas e Informática de la Universidad Nacional Mayor de San Marcos, al cual pertenece el autor. Corresponde a la actividad 3 de la metodología DSR.

En el capítulo 6, se diseña y desarrolla los modelos del artefacto. Se realizan diversos experimentos, con el conjunto de datos construido en el Capítulo 5, para determinar que el artefacto cumplía con resolver el objetivo planteado, se evalúa el rendimiento y se validan con un conjunto de datos externo para demostrar la generalización del modelo.

También se discuten y analizan los resultados obtenidos luego de la validación del artefacto construido. Corresponden las actividades 3, 4 y 5 de la metodología DSR.

1.9. Producción científica

Se culmina con el capítulo 6, donde se exponen las conclusiones de la investigación y los trabajos futuros que se identificaron de este trabajo. Esta investigación dio origen a 03 artículos científicos:

1° “Aprendizaje automático para el diagnóstico de células cancerosas en imágenes citológicas de líquido pleural: Una revisión sistemática de la literatura”, publicado en diciembre 2024 en la revista RISTI (Ver ANEXO 4).

2° “Construcción de un nuevo conjunto de datos de imágenes citológicas de Líquido Pleural para Investigación”, artículo Q2 y publicado en la revista IJOE eISSN: 2626-8493. Vol. 21 No. 7 (2025) con doi.org/10.3991/ijoe.v21i07.54323 de acceso abierto. (Ver ANEXO 5).

3° “Clasificación de imágenes de citología pleural mediante aprendizaje profundo y votación por conjunto de los modelos ResNet, DenseNet e Inception”, ha sido enviado a la revista EXPERT SYSTEMS WITH APPLICATIONS para una publicación Q1 (Ver ANEXO 6).

Esta investigación ha recibido financiamiento de la Universidad Nacional Mayor de San Marcos – R. R. N.º 005446-2025-R y Número de proyecto C25200641 – Tipo de proyecto y año PCONFIGI 2025.

Corresponden a la actividad 6 de la metodología DSR, incluyendo la presentación, aprobación, exposición y sustentación de esta investigación.

CAPÍTULO 2. MARCO TEÓRICO

El Machine Learning (ML) destaca como una herramienta clave en salud al posibilitar el análisis masivo de datos y la detección de patrones complejos que apoyan el diagnóstico médico. En el área de la citología se puede utilizar para diferenciar entre células benignas y células malignas, los métodos tradicionales presentan limitaciones por la complejidad celular, la variabilidad de las muestras y la necesidad de especialistas experimentados.

Los modelos de ML, y en particular las redes neuronales convolucionales (CNN), permiten automatizar procesos como el preprocesamiento de imágenes, la segmentación de núcleos y la clasificación celular, alcanzando altos niveles de sensibilidad y especificidad. Estas herramientas no buscan reemplazar al especialista, sino complementar su labor, reduciendo el tiempo de análisis, estandarizando procesos y mejorando la precisión diagnóstica.

En este marco, el uso de ML en citología pleural representa un paso hacia una medicina más precisa y personalizada, con potencial para integrarse a datos clínicos y moleculares, ampliando sus aplicaciones en la detección de malignidad y en la orientación terapéutica.

Por tanto, es necesario revisar los trabajos existentes y abordar la IA como una herramienta a utilizar. Entre los temas a estudiar destacan las técnicas médicas, las técnicas de diseño e implementación de los modelos inteligentes, la necesidad de crear un conjunto de datos comunes en el país.

2.1. Marco filosófico o epistemológico

El marco filosófico o epistemológico de esta tesis se sustenta en el positivismo científico y el realismo crítico. La combinación de estos dos enfoques permite tomar el problema de investigación de forma integral ya que guían el desarrollo metodológico al priorizar la verificación empírica y la explicación objetiva de los fenómenos observados en el ámbito médico y computacional. El Positivismo, proporciona una base para recolectar y analizar

datos objetivos; y el Realismo crítico, facilita la comprensión de las causas subyacentes y las interrelaciones en los datos, lo que es esencial para la interpretación clínica de los resultados.

El Positivismo Científico afirma que el conocimiento debe basarse en hechos observables y medibles. En este contexto, las imágenes citológicas digitalizadas se consideran evidencia objetiva que puede analizarse mediante herramientas computacionales como las redes neuronales convolucionales (CNN). Este enfoque permite generar resultados replicables y cuantificables, lo que resulta crucial para validar la efectividad de los modelos propuestos.

El Realismo Crítico complementa al positivismo al reconocer que los fenómenos observados, como el comportamiento de las células en el líquido pleural, están influenciados por estructuras subyacentes que no siempre son directamente observables. En este caso, las CNN actúan como herramientas para desentrañar estos patrones ocultos, identificando relaciones complejas entre las células malignas y su entorno.

La convergencia de filosofías, en el contexto del aprendizaje profundo, se justifican con la elección de modelos basados en datos (positivismo) y la exploración de sus implicancias clínicas (realismo crítico). Además, el aprendizaje por transferencia y el uso de clasificadores combinados como el voting classifier refuerzan esta convergencia al aprovechar conocimientos preexistentes para resolver problemas complejos en nuevos dominios.

2.2. Antecedentes del problema

El objetivo de esta sección es situar el problema de investigación en un contexto de revisión histórico y teórico. Los artículos seleccionados fueron analizados para conocer cómo ha avanzado el conocimiento en el tema, especialmente se enfocaron en las investigaciones sobre el problema de investigación y como se ha venido trabajando.

El derrame pleural maligno (DPM) es definido en (Tetik Kurt et al., 2018) como la presencia de células malignas en el líquido pleural. En ese trabajo, también mencionan,

que puede detectarse con biopsias, toracoscopia y autopsia. Para (Araujo, 2020) el DPM es un problema clínico común en pacientes con enfermedad neoplásica, cuando ocurre es el significado de una enfermedad neoplásica metastásica, que a veces se detecta en un estado muy avanzado que podría significar un corto tiempo de vida para el paciente. En (Ferreiro et al., 2020) relatan que en exámenes post mortem se encontró que el DPM estuvo en el 15% de los pacientes que fallecieron con enfermedad maligna. El diagnóstico del DPM, que se asocia a una baja esperanza de vida, se confirma por la presencia de células malignas en el líquido y/o en el tejido pleural. En ese contexto, Budak et al. (2023) indican que la citología puede ser útil para fines diagnósticos y pronósticos. En (Huang et al., 2020) consideran que el diagnóstico de células cancerosas en líquido pleural es fundamental para determinar el manejo clínico, pero requiere recursos tecnológicos avanzados y personal capacitado.

Estudios internacionales destacan que el examen citológico de líquido pleural tiene una sensibilidad promedio del 60%, pero este porcentaje puede mejorar significativamente con la implementación de métodos automatizados basados en inteligencia artificial. Por ejemplo, en (Zhou et al., 2021) resaltan el avance tecnológico en países desarrollados que ha permitido la incorporación de sistemas basados en Machine Learning para analizar imágenes de muestras de líquido pleural, mejorando la precisión del diagnóstico y reduciendo la dependencia de especialistas. Por otro lado, (Vásquez et al., 2023) enfatiza que esta tecnología aún no ha sido ampliamente implementada en el Perú, donde las provincias enfrentan barreras económicas, logísticas y de capacitación. Pero, lo cierto es que la tecnología es para todos, es universal, a no ser que sea patentada, en particular en asuntos médicos, como dice (Guerra & García-Mayor, 2018): “Es evidente que la inteligencia artificial puede ayudar a los médicos a realizar un diagnóstico al instante, principalmente en especialidades en las que se requiere el escrutinio de muchas imágenes, como la radiología o la anatomía patológica. Y eso es incluso más importante en áreas rurales o en países en vías de desarrollo, donde apenas hay especialistas. Es, también, evidente que el proceso todavía laborioso y costoso, se mejorará, y que en un futuro próximo los médicos podrán usar la IA para tomar mejores decisiones y proporcionar el mejor cuidado posible a los pacientes”. Para (Huertas et al., 2022), los centros de salud en

las provincias del Perú suelen carecer de equipos de laboratorio idóneos, lo que obliga a los pacientes a desplazarse largas distancias hacia la capital o a centros regionales, incrementando los costos y los tiempos de espera. Como Obermeyer et al., (2017), manifiesta que al inicio de la práctica médica los doctores se reunían y discutían los casos clínicos en forma conjunta, tratando de resolver casos complejos que individualmente no podrían lograrlo ya que se dispone de varias aristas que sobrepasan la capacidad para comprender. En consecuencia, como enfatiza (Gui & Chan, 2017), surge la necesidad de encontrar herramientas que puedan incorporar gran cantidad de datos, identificar patrones y elaborar modelos de ayuda a las limitaciones de los humanos, bajar el trabajo médico, rapidez en la atención, otorgar servicios personalizados y amenorar costes.

2.3. Bases Teóricas

A continuación, se presenta los conceptos que se han utilizado en el desarrollo de este trabajo y que dan soporte a la investigación. Siendo el tema central los modelos computacionales. También la Citología, Derrame Pleural, Examen citológico, Aprendizaje Profundo, Redes Neuronales Convolucionales, Aprendizaje por Transferencia, Arquitecturas: ResNet50V2, DesNet e Inception y Voting clasifier.

2.3.1. Citología

El diagnóstico de muchas enfermedades, en particular el cáncer, requiere una evaluación microscópica de una muestra de fluido corporal. La citología es el campo que utiliza la muestra de tejido más pequeña posible para el diagnóstico, reduciendo los criterios tanto como sea posible al nivel celular (Fisher., 2014).

La citología es vital en la medicina, ofrece capacidad de diagnóstico sin medidas invasivas y rápidas para el cáncer, infecciones y el análisis celular la convierten en una herramienta fundamental en la práctica médica moderna. Ofrece alta precisión, simplicidad y bajo costo, lo que la convierte en un método diagnostico atractivo con mínimas complicaciones.

La citología se utiliza en el diagnóstico y el pronóstico del cáncer, determinando la presencia y la extensión de la enfermedad neoplásica y evaluando la clasificación y estadificación del tumor (Varone & Troncone, 2012). Es un procedimiento diagnóstico indispensable para la detección temprana, el diagnóstico y el seguimiento de diversas afecciones en los órganos del cuerpo. En la Figura 3 se observa las aplicaciones de la citología.

La revolución digital está en marcha y se espera que las nuevas tecnologías afecten la práctica de la citología. Los avances en la investigación incluyen el desarrollo de métodos de imágenes digitales para adquirir y reproducir imágenes digitales totalmente enfocadas de preparación citológicas, mejorando la precisión y la velocidad del diagnóstico.

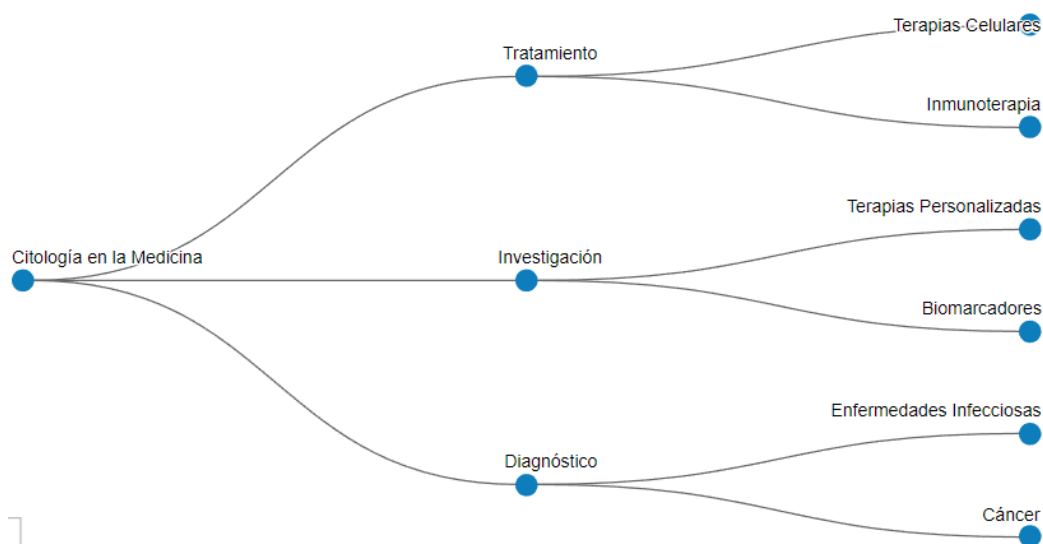


Figura 3. Aplicaciones de la citología

Fuente: Tomada de <https://www.scopus.com/search/form.uri#scopus-ai>

En la Tabla 1, se muestra las diferentes aplicaciones que existen para la citología en la medicina.

Tabla 1. Aplicaciones de la citología en la medicina

Aplicaciones de la citología	Resumen
Examen citológico	Para diagnostico en muestras ginecológicas, no ginecológicas, fluidos corporales y aspirados de aguja fina
Enfermedades infecciosas	Identifica bacterias, hongos, virus y parásitos
Dermatología	Identifica bacterias, levaduras, fúngicos, parásitos, células inflamatorias y morfología de queratinocitos
Tele citología	Consulta remota sobre muestras citológicas mediante imágenes digitales, facilitando la colaboración y el diagnóstico a distancia
Citoquímica e inmunohistoquímica	Para mejorar la calidad de las preparaciones de las muestras y para detectar marcadores de cáncer

Fuente. Elaboración propia.

2.3.2. Examen citológico

La citología de líquido pleural se distingue por ser una técnica no invasiva pero altamente informativa en la evaluación de la salud pulmonar. A través de la inserción de una fina aguja en la cavidad pleural, se obtiene una muestra del líquido que rodea los pulmones, permitiendo así el análisis microscópico de las células presentes en él. Los especialistas patólogos, procesan y examinan la muestra para identificar células anormales, tales como células cancerosas, células mesoteliales, y otras indicativas de infecciones o procesos inflamatorios. Este procedimiento no solo es valioso por su capacidad para detectar patologías subyacentes, sino también por su relativa simplicidad y su baja tasa de complicaciones. La citología del líquido pleural es el examen más simple que permite confirmar la malignidad del derrame, para el diagnóstico de DPM se debe demostrar la presencia de células malignas en el líquido pleural. La citología pleural es la prueba menos invasiva para el diagnóstico de DPM (El-Sheikh, 2012). Es una prueba diagnóstica más sensible que la biopsia pleural; ambas pruebas son complementarias, y una de ellas puede

ser positiva mientras que la otra sea negativa, aumentando la sensibilidad diagnóstica con la combinación de ambas. La sensibilidad diagnóstica de la citología del líquido pleural muestra gran variación (62 al 90%) y para biopsia pleural varía del 40 al 75% (Recuero et al., n.d.). Muchos de los DPM se deben a metástasis en la pleura, generalmente de pulmón y cáncer de mama. La toracocentesis y la biopsia pleural son dos de las técnicas más utilizadas en el estudio del DP de etiología desconocida (Araujo, 2020).

2.3.3. Derrame *Pleural*

El Derrame Pleural (DP), es una acumulación de líquido dentro del espacio pleural. Tiene múltiples causas y en general se clasifican como trasudados o exudados. La detección es por examen físico y radiografía de tórax; para determinar la causa, se requiere la toracocentesis y el análisis del líquido pleural.

Según Ferreiro et al. (2020), el DP es una patología frecuente en la práctica clínica. Afecta aproximadamente a 1,5 millones de personas en Norte América cada año. Dentro de las causas más frecuentes se destaca la falla cardíaca congestiva, el derrame para neumónico, la malignidad y la embolia pulmonar (Porcel, 2018). Así mismo, (Porcel et al., 2014) manifiesta que esta situación cambia en el continente americano, donde se menciona la tuberculosis, neumonía y la malignidad como las principales causas del derrame pleural. A su vez (Shen et al., 2017) manifiestan que la mortalidad es alta (10-20 %) si el derrame es de etiología infecciosa, donde el diagnóstico y tratamiento (antibiótico según la epidemiología local y los factores de riesgo para la resistencia bacteriana, además del drenaje según el caso y la preferencia del especialista) impactará en el pronóstico y supervivencia de los pacientes.

La prevalencia estimada de derrame pleural es de 320 casos por cada 100,000 personas en países industrializados, con una distribución de etiologías relacionadas con la prevalencia de enfermedades subyacentes, (Sahn, 2006).

Según MD.Saude (2023), el derrame pleural no es una enfermedad, sino una manifestación común de diferentes enfermedades. Por consiguiente, una vez diagnosticada la presencia de derrame pleural, el siguiente paso es investigar su causa. En

la Figura 4, observamos un pulmón con características normales y el otro con derrame pleural.



Figura 4. Pulmón normal y pulmón con Derrame Pleural

Fuente. https://www.farmaceuticos.com/pam_revista/427/page/4/

El derrame pleural se clasifica en transudado o exudado en base a los criterios de Ligth (Light, 2002). Es muy importante la diferencia ya que, en los transudados, la pleura está sana y la patología que provoca el derrame suele ser sistémica. En los exudados, la pleura suele estar enferma, tomado de (Oyonarte, 2015).

2.3.3.1 Derrame Pleural Maligno

El derrame pleural maligno (DPM) representa entre el 15 % y 35 % de los derrames pleurales y afecta al 15% de los pacientes con diagnóstico oncológico (Ferreiro et al., 2020). Es importante considerar que puede haber afectación neoplásica de la pleura sin la presencia de derrame. La mediana de supervivencia varía entre 4 y 6 meses. Generalmente los DPM son consecuencia de metástasis pleurales, que se originan en tumores pulmonares y cáncer de mama. En mujeres con diagnóstico de cáncer de mama, la

presencia de un derrame pleural ya sea en el mismo lado o en el lado opuesto al tumor primario, sugiere metástasis como primera opción.

En la Figura 5 se presenta un algoritmo para el abordaje de pacientes con sospecha de DPM.

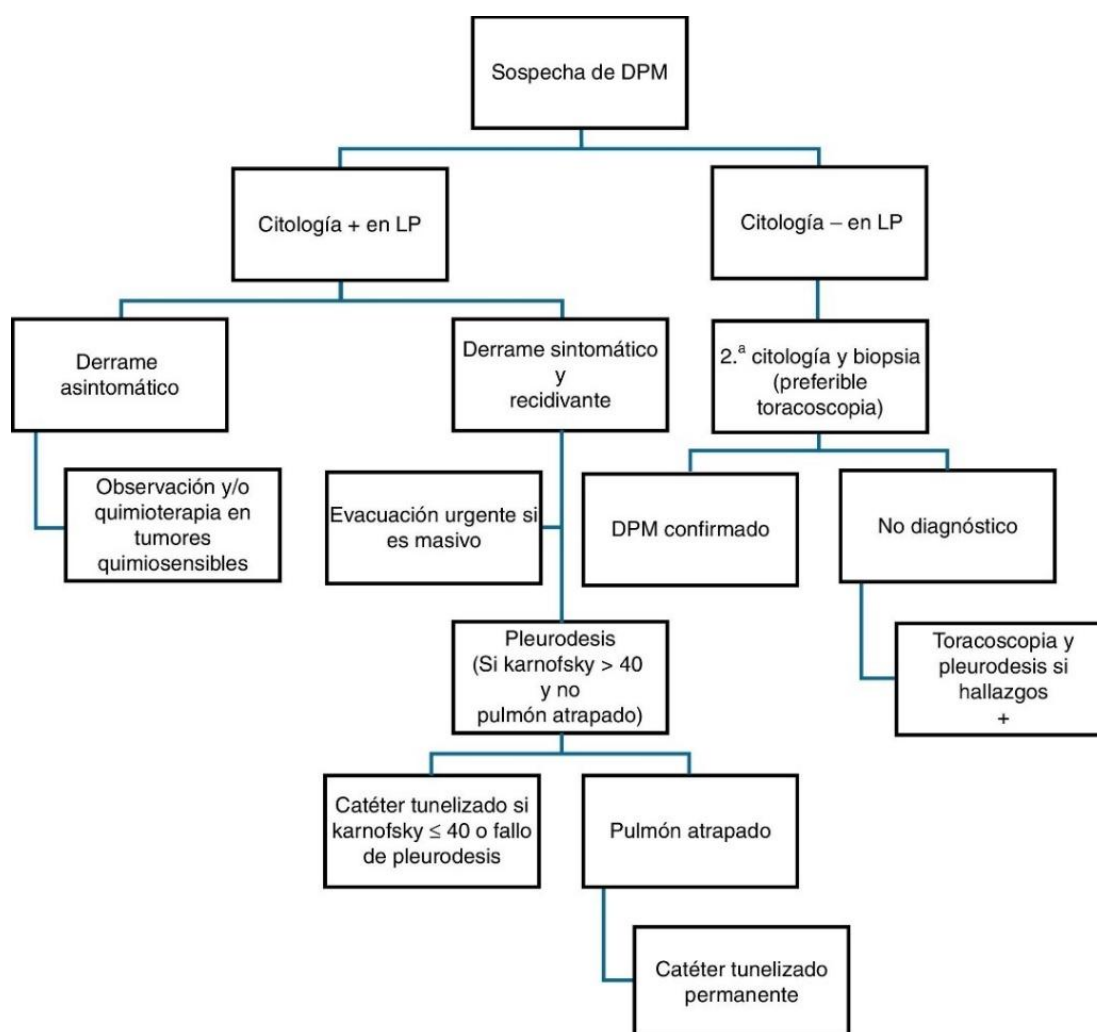


Figura 5. Sospecha de DPM.

Alrededor de 150,000 DPM se diagnostican cada año en Estados Unidos. Las causas más frecuentes son el cáncer de pulmón y de mama. Debido a que el DPM se considera una enfermedad muy avanzada, su pronóstico suele ser siempre malo (Bashour et al., 2022). El DPM se produce cuando se comprueba la presencia de células cancerosas en el líquido pleural de un paciente diagnosticado con Derrame Pleural, indica que la enfermedad está avanzando (Karadayi & Sahin, 2021).

Para diagnosticar el DPM lo convencional y más usado es el examen citológico del líquido pleural y/o una biopsia del tejido pleural, además de otros exámenes de diagnóstico e imágenes (Dixit & Agarwal, 2017). Para realizar el diagnóstico de un DPM, se precisa una citología y/o una histología a través de una biopsia pleural. Sin embargo, la historia clínica del paciente, su sintomatología y las técnicas de imagen pueden hacer sospechar su presencia. El diagnóstico acertado de malignidad en el líquido pleural puede conseguirse sólo mediante la aparición de células cancerosas en el líquido por toracocentesis y detectadas mediante una técnica denominada citología. La toracocentesis viene a ser el primer procedimiento en el algoritmo diagnóstico del derrame pleural, este puede proporcionar un diagnóstico certero cuando se sospeche malignidad (Araujo, 2020). Se encontraron varios métodos, pero de acuerdo con El-Sheik (2012), el método más sencillo, útil, económico y seguro es el examen citológico.

En la Figura 6 se describen algunos métodos de diagnóstico para el DPM.

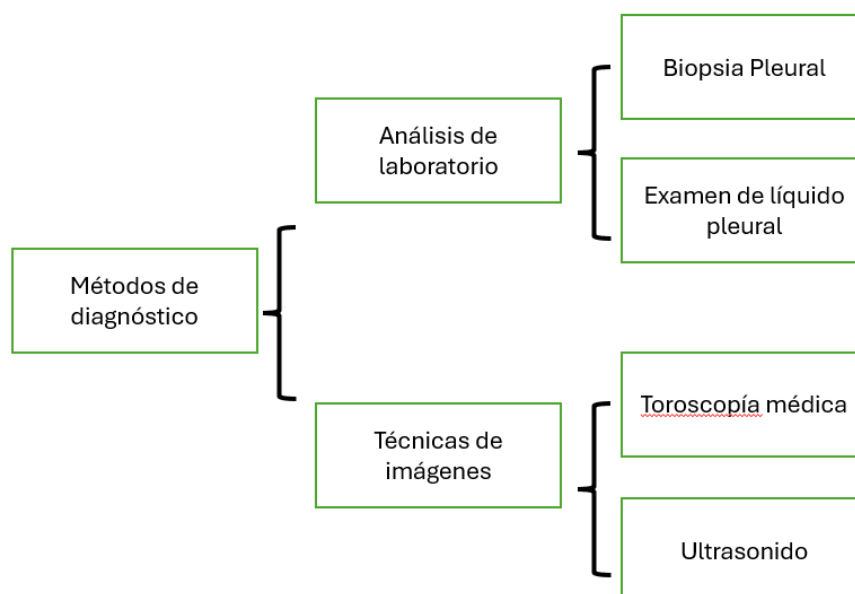


Figura 6. Métodos de diagnóstico

De la literatura revisada podemos concluir que para un diagnóstico a tiempo de DP existen diferentes procedimientos como una revisión física, examen de líquido pleural y pruebas de imágenes. Cuando se encuentran células cancerosas en el líquido pleural se llama DPM, esta enfermedad es un reto para los especialistas médicos aun cuando encontramos varias formas para su control considerando un largo plazo de vida, sin embargo, los tratamientos continúan siendo paliativos incidiendo en dar una mejor calidad de vida a los pacientes y tratar de minimizar los síntomas.

Para el diagnóstico del DPM se pueden usar muchas técnicas, pero muchos investigadores consideran fundamental el examen citológico del líquido pleural, aunque su sensibilidad en el diagnóstico del DPM es limitada, pero es de fácil acceso, no invasiva, precisa y rápida.

2.3.4. Aprendizaje profundo

El Aprendizaje Profundo (DL) es parte del aprendizaje automático, se basa en el uso de redes neuronales artificiales con múltiples capas para analizar y aprender de grandes

volúmenes de datos. Conocido como redes neuronales profundas, son capaces de identificar patrones complejos y realizar tareas como el reconocimiento de imágenes, el procesamiento del lenguaje natural y la predicción de series temporales con alta precisión (Lecun et al., 2015a).

Estas redes están inspiradas en la estructura del cerebro humano y son capaces de modelar y aprender representaciones complejas de los datos, lo que permite a los modelos identificar patrones en imágenes, texto, y otros tipos de datos (Goodfellow et al., 2016).

Su metodología implica muchas etapas, inicia con la preparación de los datos y su preprocesamiento. Luego, los datos se dividen en conjuntos de entrenamiento, validación y prueba. La red neuronal profunda se entrena mediante retro propagación, ajustando los pesos de las conexiones neuronales en función del error entre la predicción y la salida esperada. El entrenamiento es intensivo en cómputo y puede requerir grandes volúmenes de datos y tiempos de procesamiento prolongados (Lecun et al., 2015a).

Clasificación por topología de red

Las Redes Neuronales son un sistema aprendizaje y procesamiento automático inspirado en el funcionamiento del sistema nervioso humano.

Podemos observar en la Figura 7, la clasificación por topología de red

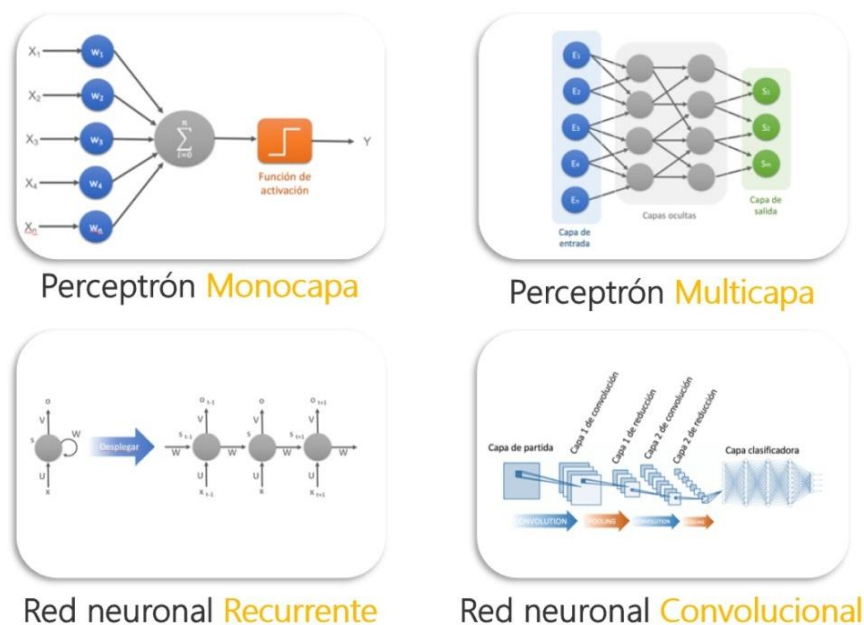


Figura 7. Cuatro topologías de red neuronal

2.3.5. Redes neuronales Convolucionales

Las redes neuronales convolucionales (CNN) son una variante del perceptrón multicapa, con la diferencia que las CNN realizan procesos de convolución entre los parámetros de la red neuronal. Se inició con la primera red neuronal convolutiva creada por Yann Le Cun en el año 1989, la cual consistió en el reconocimiento de letras manuscritas, fue creada con varias capas que implementan la extracción de características y luego la clasificación de este. Estos datos de entrada a la capa alimentan luego las demás capas para así detectar líneas verticales, vértices, entre otras (Bishop, 2024).

Las CNN pertenecen a la familia del Aprendizaje Profundo (DL) y son diferentes algoritmos diseñados para análisis de datos principalmente en imágenes. Se usan en distintas áreas como la visión por computadora, procesamiento del lenguaje natural e imágenes médicas (Xue, 2021). Las CNN es una forma de red con estructura profunda y cálculos de convoluciones, pueden extraer características específicas de las imágenes lo que las convierten en herramientas eficaces para diagnosticar malignidad o cáncer en

imágenes. Las CNN son como expertas en aprender, tomar características y analizar imágenes médicas. Pueden identificar características para clasificar tumores benignos y malignos, lo que favorece al diagnóstico preciso para clasificar los tumores en la práctica médica. Las CNN necesitan un gran tamaño de datos para identificar las características más importantes, esto permite tomar decisiones basadas en datos y la detección a tiempo de la malignidad. Las CNN se utilizan para categorizar imágenes, analizar contenidos de videos, reconocimiento de lenguajes de señas y diagnósticos en imágenes médicas (Deng, 2022).

A diferencia de las arquitecturas tradicionales basadas en el perceptrón, las CNN incluyen una o más capas convolucionales que extraen características de las imágenes de entrada a través de operaciones de convolución. Una imagen digital puede ser definida como una función bidimensional o tridimensional discreta. Está compuesta por un número finito de elementos, denominados píxeles, cada uno de los cuales tiene una ubicación y un valor determinados (Gonzales & Wood, 2018). Las CNN son capaces de capturar con éxito las variaciones espaciales en una imagen mediante la aplicación de filtros de convolución.

Estructura de una CNN

La estructura de una CNN típica está conformada principalmente por las siguientes capas como se muestra en la Figura 8, puede dividirse en dos fases: la fase de extracción de características y la fase de clasificación (Zeiler & Fergus, 2014).

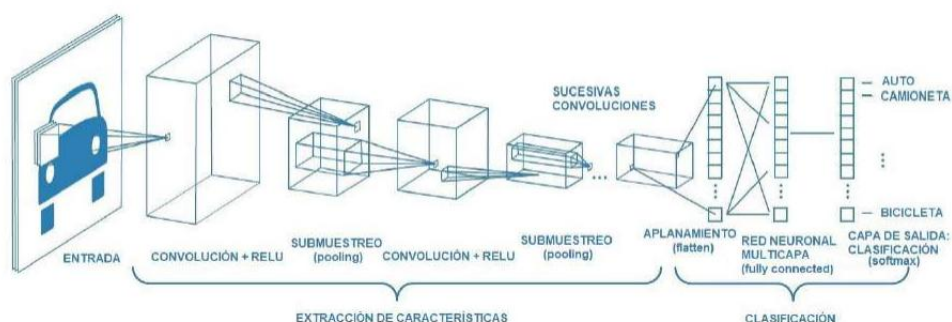


Figura 8. Estructura de una CNN.
Fuente. Tomado de (Zeiler & Fergus, 2014).

Capa de Convolución

La convolución es fundamental para la extracción de características y la reducción de la cantidad de parámetros en una CNN. Esta capa le da nombre a la red y es el elemento central. La operación de convolución consiste en aplicar un filtro o kernel sobre la imagen de entrada, generando un mapa de características que describen las particularidades de las imágenes de entrada como se muestra en la Figura 9. Esto permite a la red capturar patrones espaciales y jerárquicos en los datos visuales (Krizhevsky et al., 2017).

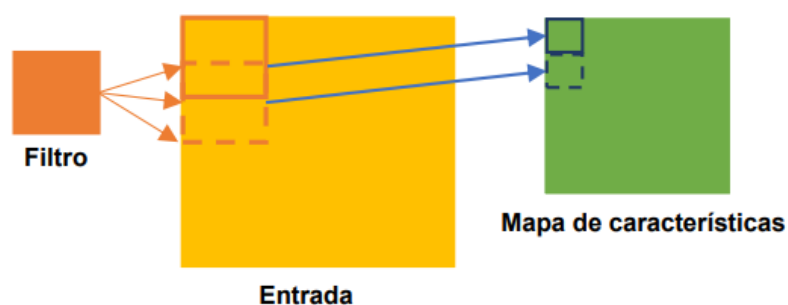


Figura 9. Capa de convolución (filtro, entrada) y mapa de características

Capa de activación.

Una vez creado el mapa de características, se pasa cada valor del mapa a través de una función de activación que aporta no linealidad, de forma muy parecida a lo que se hace con las salidas de una capa totalmente conectada. La función más utilizada es la función ReLU (Rectified Linear Unit) que, además de permitir mapeos no lineales, es matemáticamente simple y computacionalmente eficiente.

Capa de submuestreo o pooling.

El max-pooling es una operación de muestreo utilizada en las CNN para reducir la dimensionalidad de los mapas de características. Selecciona el valor máximo en una región determinada de la imagen, preservando las características más importantes y reduciendo la sensibilidad a pequeños desplazamientos y variaciones (Goodfellow et al., 2016).

Una limitación del enfoque convolucional es que las características descubiertas por el filtro dependen fuertemente de la localización espacial de los objetos o texturas presentes. Razón por la cual, la rotación, el desplazamiento u otros cambios menores en la imagen de entrada darán como resultado un mapa de características diferente, cuando es deseable que las características descubiertas sean independientes de estos cambios. Para abordar este problema se utiliza la capa de pooling, en la cual los valores resultantes del filtrado son submuestreos, reduciendo el tamaño final de las matrices. En la Figura 10 se muestran los ejemplos de pooling con un tamaño 2x2. En max-pooling la matriz resultante se obtiene como el máximo de los valores contenidos en regiones de 2 por 2 píxeles. En el caso del averagepooling se reemplazan por el promedio. La idea del pooling es cubrir una región mayor del espacio, ampliando el campo de “aprendizaje” de la red haciéndolo más general.

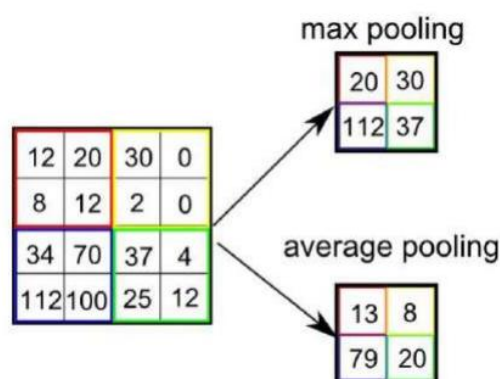


Figura 10. Matriz de entrada con max-pooling y average-pooling

Fuente. Tomado de (Goodfellow et al., 2016).

Capa totalmente conectada o fully connected.

Las capas completamente conectadas transforman los mapas de características en una salida final, comúnmente utilizada para la clasificación (LeCun et al., 2015).

Típicamente se alternan capas de conexión, ya que es una forma de aprender combinaciones no lineales de las características de alto nivel representadas por la salida de la capa convolucional. Además de las capas anteriores, las CNN pueden incluir capas con perceptrones.

Fase de extracción de características, alterna varias capas de neuronas convolucionales, capas de activación y capas de submuestreo (pooling) que reducen subsecuentemente la cantidad de información (píxeles). Como resultado, la salida de esta fase puede interpretarse como un conjunto de características asociadas a la imagen ingresada. El apilamiento de capas convolucionales en esta fase es lo que permite realizar una descomposición jerárquica de la entrada y que las capas más profundas del modelo aprendan características más abstractas (complejas).

Fase de clasificación es una red neuronal tradicional, del tipo perceptrón multicapa, cuya arquitectura se ajusta según la clasificación a realizar. La cantidad de neuronas de salida iguala la cantidad de clases.

En el contexto del aprendizaje profundo la mayoría de los modelos que resuelven problemas complejos necesitan una gran cantidad de datos con el fin de optimizar la gran cantidad de parámetros que estos modelos poseen. Para los modelos supervisados, obtener grandes conjuntos de datos etiquetados puede ser realmente difícil, considerando el tiempo y el esfuerzo que se necesita para el etiquetado. Un ejemplo de ello sería ImageNet (Deng et al., 2009), una base de datos a gran escala que se compone de millones de imágenes pertenecientes a más de 20.000 categorías diferentes, fuertemente inspirada en la base de la estructura WordNet, y cuya conformación llevó muchos años de etiquetado manual. En la Figura 11 se muestra imágenes al azar de ImageNet.

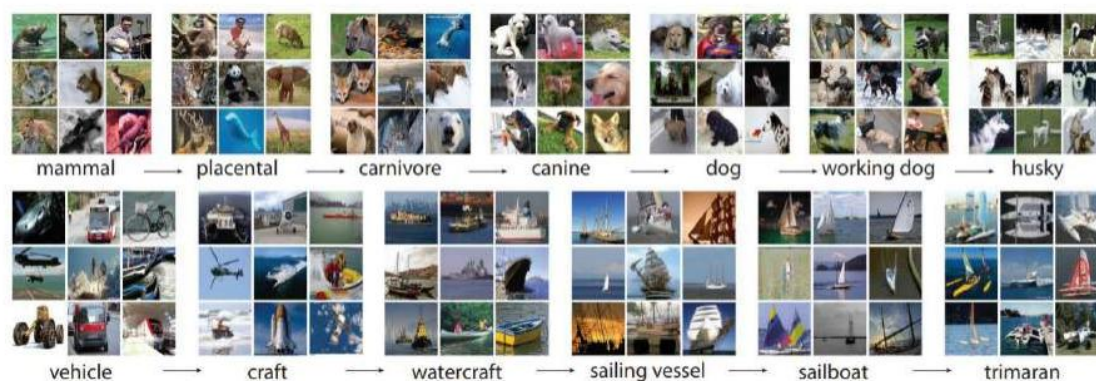


Figura 11. Imágenes al azar de Imagenet

Fuente. Imagenet.

Las principales arquitecturas de CNN surgieron del ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), un concurso de visión computacional desarrollado anualmente desde 2010 y hasta 2017, a partir de subconjuntos de datos obtenidos de ImageNet. Algunas de las arquitecturas pre-entrenadas que surgieron en el marco de este concurso son: AlexNet (Krizhevsky et al., 2017), VGGNet (Simonyan & Zisserman, 2014), Inception (GoogLeNet) (Szegedy et al., 2014), ResNet (He et al., 2015), DenseNet (G. Huang et al., 2017), entre otras.

2.3.6. Arquitectura ResNet50V2

Fue introducida por (He et al., 2015), en su artículo "Deep Residual Learning for Image Recognition" en 2015. La arquitectura ResNet50V2 es una red residual profunda, permite que el entrenamiento sea estable mediante conexiones residuales y una mejor organización de las operaciones, obteniendo gran desempeño en labores de visión por computadora. La arquitectura esta formada por 50 capas profundas, estructuradas en bloques residuales, combinación de capas convolucionales y atajos. Los bloques residuales son la clave para resolver el problema del desvanecimiento del gradiente, al permitir que el flujo de información pase sin modificaciones a través de los atajos (saltar conexiones). Cada bloque residual tiene un atajo que salta sobre varias capas, que facilita a evitar la

degradación del desempeño en redes profundas. Aplica normalización por lotes antes de la activación (ReLU) y la convolución para optimizar el flujo de gradientes.

El flujo de datos por medio de las convoluciones iniciales en la entrada pasa por una capa convolucional inicial seguida de normalización y agrupación (pooling). Tiene 4 bloques principales, con el número de filtros aumentando progresivamente (64, 128, 256, 512). Las capas finales terminan con un promedio global de pooling y una capa completamente conectada (fully connected) para la clasificación.

Características principales:

Conexiones residuales: ResNet50 utiliza conexiones residuales para facilitar el entrenamiento de redes neuronales profundas. Estas conexiones residuales permiten que la información original fluya directamente a través de la red sin ser transformada, lo que ayuda a mitigar el problema del desvanecimiento del gradiente y facilita el entrenamiento de redes profundas.

Bloques residuales: ResNet50 está compuesto por bloques residuales que contienen varias capas convolucionales seguidas de una conexión residual. Estos bloques residuales se repiten varias veces a lo largo de la red para construir una arquitectura profunda.

Arquitectura en cascada: ResNet50 sigue una arquitectura en cascada donde las capas se organizan en bloques, y cada bloque contiene múltiples capas convolucionales seguidas de una conexión residual. La arquitectura en cascada permite construir redes neuronales muy profundas con miles de capas, lo que ha demostrado ser beneficioso para tareas de visión por computadora.

Capas de agrupamiento (pooling): ResNet50 también incluye capas de agrupamiento (pooling) para reducir la dimensionalidad de las características y aumentar la invariancia a la translación de estas. Estas capas de agrupamiento ayudan a mejorar la eficiencia computacional y a evitar el sobreajuste.

Pre-entrenamiento y transferencia de aprendizaje: Al igual que otras arquitecturas ResNet, ResNet50 se puede utilizar como una red pre entrenada para transferir aprendizaje a otras tareas o conjuntos de datos específicos.

Por lo que el modelo se puede entrenar inicialmente en un conjunto de datos grande (como ImageNet) y luego ajustar o fine-tunear para una tarea específica con un conjunto de datos más pequeño.

¿Qué es el bloque residual?

Los bloques residuales se consideran el componente básico de ResNet, ver Figura 12. En el bloque residual, la entrada x se agrega directamente a la salida de la red, es decir, $F(x) + x$ y esta ruta se conoce como conexión de salto o acceso directo.

Aquí la salida será $H(x) = F(x) + x$.

Entonces, cuando $F(x) = 0$, entonces $H(x) = x$, lo que se conoce como mapeo de identidad. es decir, cuando la entrada de la red es igual a su salida.

Para sumar $F(x) + x$, la forma de ambos debe ser exactamente la misma. Si la forma no es la misma, multiplicamos una matriz W_s con la entrada x .

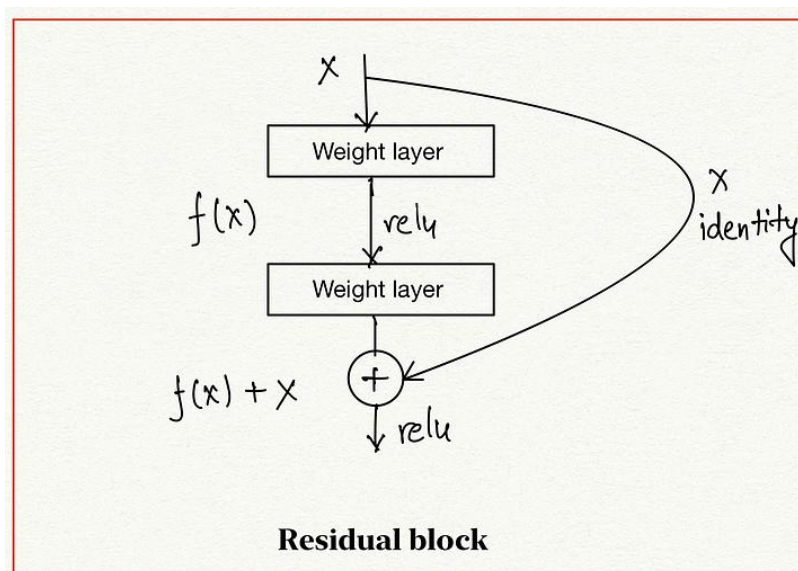


Figura 12. **Bloque residual.**

Fuente. Tomado de (Lecun et al., 2015a).

Para redes más profundas, Figura 13, usamos la estructura de cuello de botella, primero usando una convolución 1x1 para la reducción de dimensionalidad, luego una convolución 3x3 y finalmente usando la dimensionalidad 1x1 para restaurar la dimensión original.

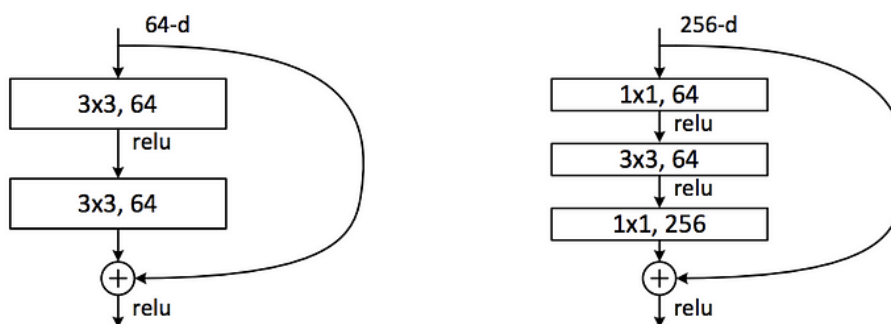


Figura 13. **Redes más profundas.**

Fuente. <https://arxiv.org/pdf/1512.03385.pdf>.

Esto equivale a reducir el número de parámetros para el mismo número de capas, por lo que se puede extender a modelos más profundos. Por lo tanto, el autor propuso ResNet con 50, 101 y 152 capas, y no solo no tenía problemas de degradación, sino que también

la tasa de error se redujo considerablemente y la complejidad computacional se volvió muy baja.

ResNet abordó el problema de optimización / degradación con una red más profunda al omitir ciertas capas utilizando conexiones de omisión o bloques residuales. Logró aumentar la profundidad y no perjudicó el rendimiento del modelo, sino que el rendimiento mejoró. También pudo reducir el número de parámetros, que fue un factor importante en términos de complejidad de cálculo.

Entre sus principales ventajas se menciona que logra un entrenamiento estable en redes más profundas gracias a los atajos y la reordenación de las operaciones (BatchNorm antes de ReLU y convoluciones) mejora el flujo de gradientes durante el entrenamiento.

2.3.7. *Arquitectura DesNet121*

Propuesta por (G. Huang et al., 2017), conecta cada capa directo con las demás capas en un bloque denso. En lugar de conexiones residuales, utiliza conexiones densas. Cada capa recibe las características de todas las capas anteriores y pasa sus propias características a las capas posteriores tal como se muestra en la Figura 14.

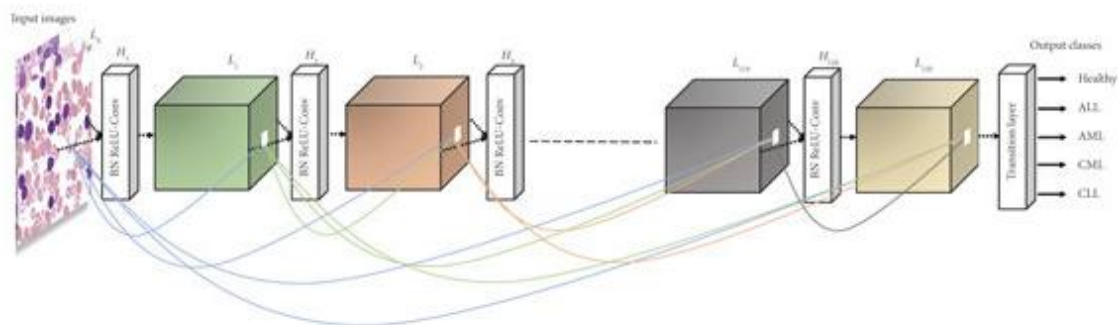


Figura 14. Arquitectura DesNet121

Fuente. (Huang et al., 2020)

Ventajas:

- Alivia el problema del gradiente de desaparición.
- Fortalece la propagación de características.
- Fomenta la reutilización de características.
- Reduce significativamente el número de parámetros.
- Buen desempeño en tareas de clasificación e imágenes

2.3.8. *Arquitectura Inception*

InceptionV3 es una arquitectura de redes neuronales profundas diseñada para tareas de clasificación de imágenes. Introduce módulos Inception que combinan convoluciones de diferentes tamaños (1x1, 3x3, 5x5) en paralelo, maximizando la eficiencia computacional y la extracción de características. Con técnicas como factorización de convoluciones y reducción de parámetros, logra un balance entre profundidad y rendimiento, siendo altamente precisa en benchmark como ImageNet. En la figura 15 se muestra InceptionV3.

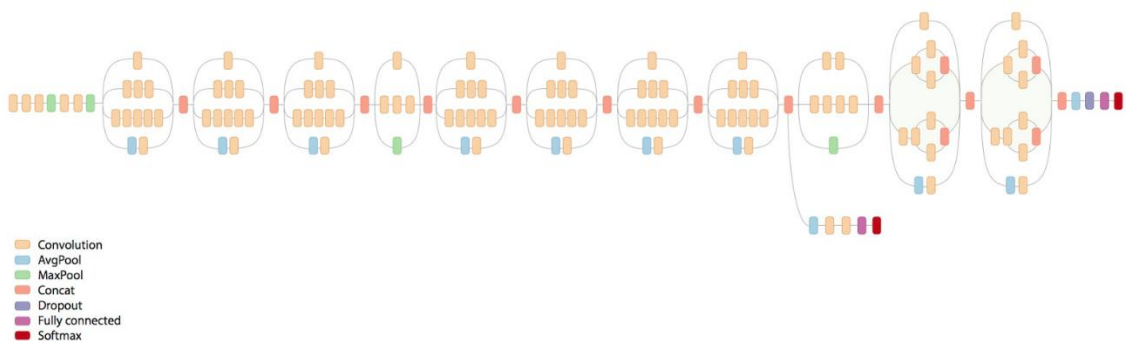


Figura 15. Arquitectura InceptionV3

Fuente. (Li & Song, 2022)

Aspectos claves de Inception:

Motivación:

Inception, también conocida como GoogLeNet, fue diseñada para mejorar la utilización de recursos computacionales dentro de la red.

El objetivo era aumentar la profundidad y el ancho de la red sin aumentar el costo computacional.

Características:

Estructura densa: Inception se basa en encontrar una estructura local óptima en una red convolucional y cubrirla con componentes densos disponibles.

Módulos Inception: Utiliza módulos con múltiples tamaños de filtros (convoluciones 1x1, 3x3 y 5x5) para capturar características a diferentes escalas.

Reducción de dimensiones: Emplea convoluciones 1x1 para reducir la dimensión antes de aplicar convoluciones más grandes.

Resultados:

Inception logró un rendimiento sobresaliente en clasificación y detección de imágenes en ILSVRC 2014.

Inception es una arquitectura que optimiza los recursos computacionales y ha dejado una huella significativa en la visión por computadora.

2.3.9. Aprendizaje *por* transferencia

Un enfoque muy utilizado en la implementación práctica de las CNN para resolver el problema de la escasez de datos etiquetados (común al trabajar con imágenes médicas) es el transfer-learning (Bhattacharya et al., 2021). Éste consiste en utilizar una CNN previamente entrenada y reemplazar la fase de clasificación por un nuevo clasificador para reentrenar la red con el conjunto de imágenes de interés, reajustando sólo los parámetros

internos de la fase de clasificación; lo que reduce drásticamente el tiempo de entrenamiento y el número de ejemplos necesarios. En la Figura 16 se muestra un esquema general del enfoque de transfer-learning.

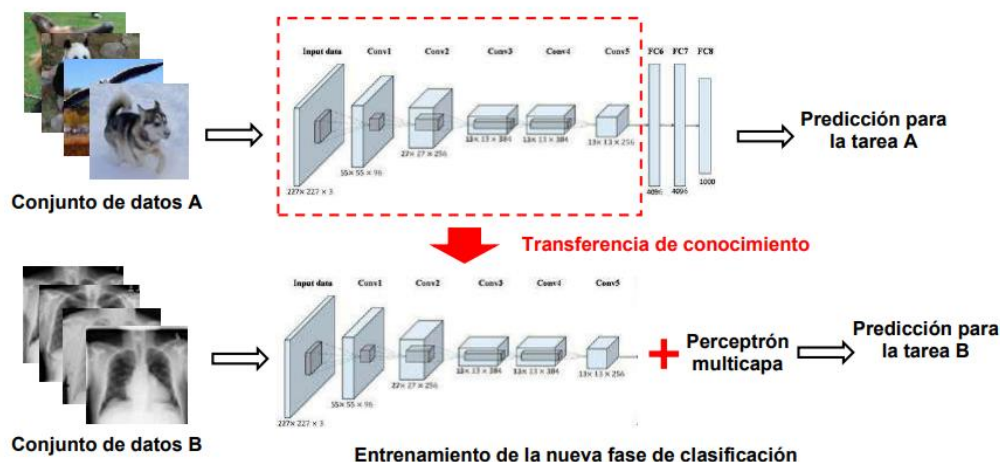


Figura 16. Transfer Learning

El uso más común del transfer-learning en el contexto del aprendizaje profundo es mediante el siguiente flujo de trabajo (Martinez & Hernandez, n.d.):

- Se toman algunas de las capas de un modelo previamente entrenado.
- Se congelan las capas del modelo pre-entrenado, para evitar perder la información que contienen en futuras rondas de entrenamiento, es decir, sus parámetros internos se fijan como no “ajustables”.
- Se añaden nuevas capas entrenables, es decir, con parámetros ajustables, sobre las capas congeladas extraídas en el paso a). Las nuevas capas aprenderán a convertir las antiguas características en predicciones sobre un nuevo conjunto de datos.
- Se entrena la red ajustando sólo las nuevas capas con el nuevo conjunto de datos. Las capas del modelo pre entrenado no se modifican. Un último paso, opcional, es

el fine-tuning, que consiste en descongelar todo el modelo ya entrenado (o parte de él), y volver a entrenarlo con los nuevos datos con una tasa de aprendizaje muy baja. De este modo se pueden conseguir mejoras significativas, adaptando de forma incremental las características pre entrenadas a los nuevos datos. Aunque sólo se toman del modelo pre-entrenado las capas de la fase de clasificación, existen diferentes enfoques de transfer-learning, algunos de los cuales sólo reemplazan y reentrenan la última capa de la fase de clasificación. En este sentido, se entiende como transfer-learning todo enfoque en el que se reutilicen capas de una red entrenada previamente. Este flujo descrito permite modificar dinámicamente los datos de entrada del nuevo modelo durante el entrenamiento, lo que permite utilizar técnicas de data-augmentation para mejorar la generalización de la red.

2.3.10. Clasificador de votación

Voting Classifier es un método de aprendizaje automático en ensamblado que combina las predicciones de varios modelos base (como árboles de decisión, SVMs o regresiones logísticas) para mejorar el rendimiento. Puede utilizar votación mayoritaria (hard voting) o promediar probabilidades (soft voting) para determinar la clase final. Es efectivo cuando los modelos individuales son diversos y tienen buenos desempeños individuales. En la Figura 17 se muestra las características del Voting Classifier.

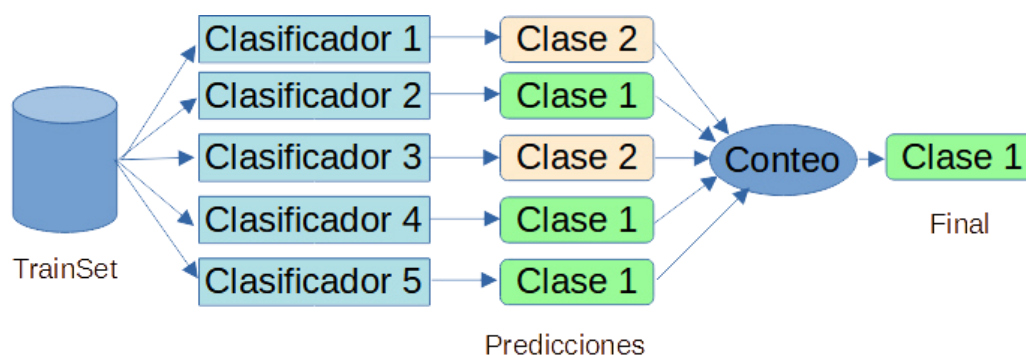


Figura 17. Voting classifier.

Situación del recurso humano MEDICO ESPECIALISTA en el MINSA

De acuerdo con la información del Registro Nacional de Personal de Salud – INFORHUS, que se muestra en la Figura 18 se tienen algunos puntos relevantes:

Oferta Total de Médicos Especialistas en el MINSA a Nivel Nacional:

El MINSA cuenta con un total de 6074 médicos especialistas. De ellos, el 61.5% pertenece a especialidades clínicas, el 33.2% a especialidades quirúrgicas, el 4.9% a especialidades de apoyo al diagnóstico y tratamiento, y el 0.4% a especialidades de salud pública.

Evolución Cuantitativa de Profesionales de la Salud en el MINSA y Gobiernos Regionales (2013-2021):

Se ha observado un crecimiento en la cantidad de profesionales de la salud en el MINSA y los Gobiernos Regionales durante este período.

Disparidades Regionales y Necesidades de Dotación:

La Dirección de Recursos Humanos del MINSA ha destacado las disparidades regionales y las necesidades de dotación en diversas áreas del país.

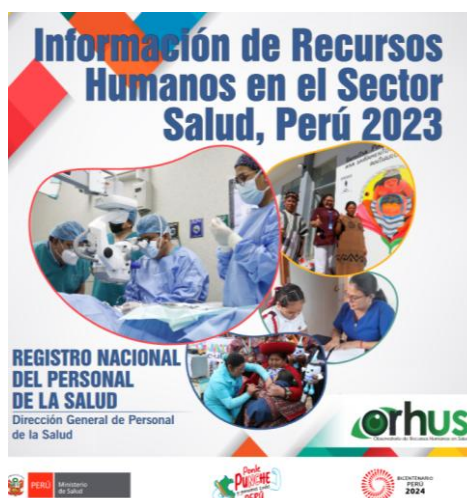


Figura 18. Informe RRHH del sector salud al 2023 publicado en 2024.

Fuente. Dirección General del Personal de la Salud.

CAPÍTULO 3. ESTADO DEL ARTE

La Revisión Sistemática de la Literatura (RSL) constituye el núcleo del aporte teórico de esta investigación, siguiendo la estructura propuesta por Hernández Sampieri (2022). Esta revisión no solo organiza el conocimiento disponible, sino que fundamenta la selección de técnicas y metodologías empleadas en esta tesis doctoral, cumpliendo la función de marco teórico-analítico.

En los últimos años, la incursión de la Inteligencia Artificial en la Patología es un campo de investigación relativamente nuevo que enfrenta muchos desafíos que se pudieron observar en la revisión de antecedentes donde se encontró estudios para diagnosticar diferentes tipos de neoplasias, sin embargo, para diagnosticar células cancerosas en líquido pleural hay mucho por trabajar (Niazi et al., 2019). En el futuro será posible la aplicación del aprendizaje automático al análisis bioquímico del líquido pleural, el diagnóstico molecular y las tecnologías de alto rendimiento (Zheng & Hu, 2022).

Existen varios enfoques relacionados con los tópicos abordados con el objetivo de este trabajo. Esas fuentes enfatizan la detección de células cancerosas en imágenes citológicas usando tecnologías computacionales, como citología computacional con técnicas de inteligencia artificial de la familia de aprendizaje automático (machine learning). Con ese objetivo, en este capítulo, se explican las fuentes bibliográficas de categoría científica de bases de datos de referencia mundial, siguiendo una metodología de revisión de la literatura.

3.1. Selección de artículos de trabajos relacionados

Para la realización de la revisión de la literatura, sobre los trabajos de modelos de diagnóstico médico en citología de líquido pleural, se definió un escenario general de investigación cualitativa para satisfacer algunas preguntas de investigación (PI) vinculadas al objetivo de este trabajo, en el contexto de técnicas y datos usados:

PI1: ¿Qué estudios de investigación de han realizado usando técnicas de Machine Learning para diagnóstico de células neoplásicas en imágenes citológicas de líquido pleural?

PI2: ¿Qué conjuntos de datos de imágenes citológicas se encuentran disponibles para el diagnóstico de células neoplásicas?

Se utilizó la estrategia PICO (Patient, Intervention, Comparison, Outcome) (Richardson et al., 1995), que representa la relación con la población, la intervención, la comparación y el resultado. La población fue determinada por los informes citológicos de líquido pleural con tinción Papanicolau con sus respectivas láminas entre enero 2023 a octubre 2024, del Servicio de Anatomía Patológica del HNCH. La intervención fueron los modelos de Aprendizaje Profundo, comparando las mejores técnicas y sus métricas tal como observamos en la Tabla 2.

Tabla 2. Estrategia PICO

Elementos	
P	Informes citológicos de líquido pleural con tinción Papanicolau con su lámina
I	Modelo basado en Aprendizaje Profundo
C	Técnicas y métricas
O	Rendimiento del modelo

También se utilizó la metodología PRISMA (Rethlefsen et al., n.d.), para identificar, seleccionar, evaluar y sintetizar estudios referentes al tema de investigación. De acuerdo con esta metodología se debe definir el objetivo, establecer criterios, seleccionar fuentes de información, planificar la búsqueda de información, seleccionar investigaciones y recopilar datos (Page et al., 2021).

El estudio analiza las técnicas existentes y propuestas para clasificar células neoplásicas en imágenes citológicas de líquido pleural. Se debe tener una comprensión general de derrame pleural, de Aprendizaje Profundo y los diversos estudios que se han propuesto o adaptado para detectarlo.

La búsqueda de las fuentes se realizó en las bases de datos electrónicas Scopus, IEEE, Medline, Web of Science, PUBMED, IEEE. Se consideró los artículos publicados del 2018 a julio 2024 en idioma inglés. Las palabras claves utilizadas para la búsqueda fueron: “Cytological images of pleural effusion”, “image diagnosis pleural effusion cancer cells”, “artificial intelligence“, “Deep Learning”, “image segmentation”. Los criterios de inclusión son: a) Aprendizaje profundo, b) citología, c) derrame pleural; d) desde 2018; los criterios de exclusión son: e) eliminar duplicados, f) eliminar estudios sin resultados de la aplicación de las técnicas mencionadas. En la Tabla 3 se observa los criterios de búsqueda que se utilizaron.

Tabla 3. Criterio de búsqueda.

Llave	Cadena de búsqueda
Cadena de búsqueda	("cytological images" OR "pleural effusion" OR "Effusion cytology" OR "computational cytology") AND ("Deep Learning" OR "Machine Learning") AND (cancer OR malignancy OR clustering OR segmentation OR nuclei)
Límites	Texto completo + imprimir; Revistas académicas (revisadas por pares).
Sinónimos	Aplicar palabras equivalentes
Modos de búsqueda	Encuentra todos mis términos de búsqueda

3.2. Desarrollo

El objetivo fue encontrar los estudios más actuales del tema. Centramos la búsqueda en las técnicas de aprendizaje profundo para diagnosticar células cancerosas. La búsqueda recopiló 1569 artículos. Se excluyeron 718 por antigüedad y se eliminaron 95 artículos después de comprobar que eran artículos terminados y artículos de revisión. En adelante los artículos completos se analizan para conocer las técnicas propuestas de diagnóstico de células cancerosas. Se evalúa la información y analiza las contribuciones a la metodología propuesta. Además, se analizan las técnicas de preprocesamiento de imágenes, extracción de características, técnicas de selección, aprendizaje profundo (DL) y Convolutional Neural Network (CNN). Se observó el uso de diferentes métodos y técnicas de extracción de características para mejorar el rendimiento y la precisión del modelo de los 67 artículos seleccionados los cuales mostraron los temas representativos en la coincidencia temática,

tipo de indicadores, data, entre otros. En la Figura 19 se observa el trabajo de selección realizado y en la Figura 20 se muestra la estrategia de búsqueda de la información.

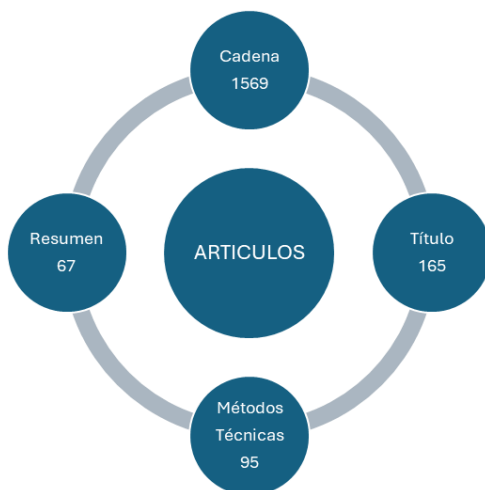


Figura 19. Selección de artículos

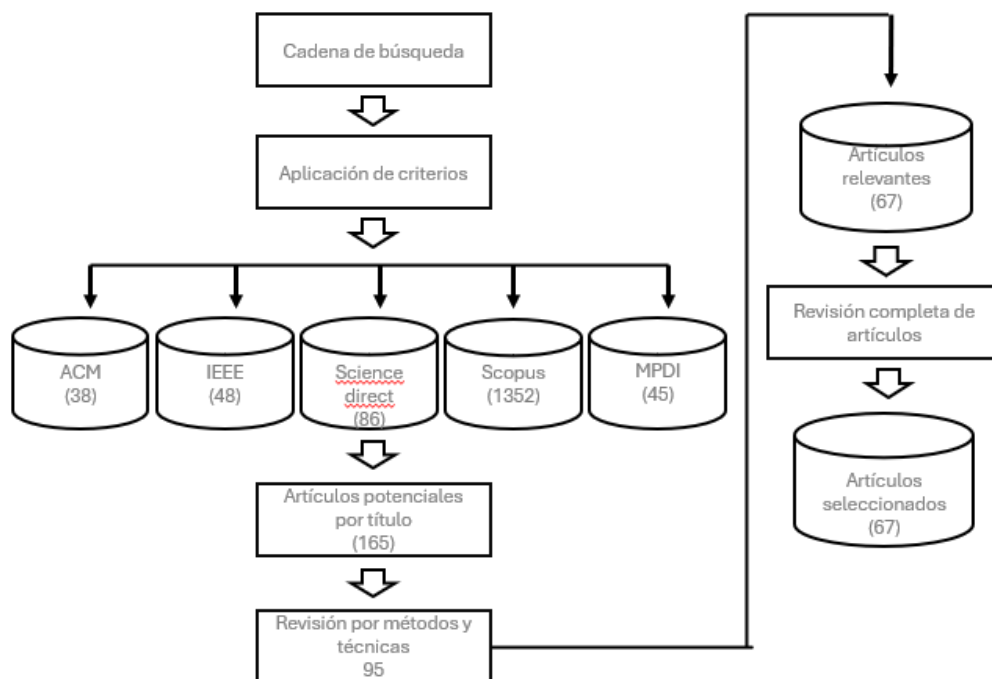


Figura 20. Estrategia de búsqueda de información

De las investigaciones encontradas, que abordan estudios sobre modelos para detectar células cancerosas en citología de derrame pleural, tal como (Lecun et al., 2015b) que destaca los avances del aprendizaje profundo en la resolución de problemas complejos en ciencia, negocios y gobierno. Este método ha demostrado ser eficiente en áreas como reconocimiento de imágenes y voz, superando a otras técnicas de aprendizaje automático en la predicción. Además, es clave para extraer estructuras complejas a partir de grandes volúmenes de datos, abriendo nuevas posibilidades en múltiples disciplinas. En (Pramesh et al., 2022) resaltan cinco prioridades para mejorar la atención del cáncer en países de bajos recursos: reducir la carga de pacientes con enfermedades avanzadas, mejorar el acceso y la asequibilidad del tratamiento, implementar atención basada en valores, optimizar la calidad de los servicios y aprovechar tecnologías innovadoras para el control del cáncer. Este enfoque promueve soluciones sostenibles en sistemas de salud limitados. En (Bera et al., 2019) analizan la eficiencia de la inteligencia artificial en patología digital para automatizar tareas repetitivas, permitiendo a los especialistas concentrarse en la toma de decisiones críticas. Este enfoque mejora la eficiencia y precisión del diagnóstico, especialmente en enfermedades con características complejas, y representa un avance importante para la práctica médica. Park, Chong, Lee, Yim, Seo, Hwang, Kim, Gong, Cho, Yoo, & Choi (2023) evaluaron un modelo DCNN basado en Inception-ResNet-V2 para detectar derrames pleurales malignos en 569 preparaciones citológicas asociados al cáncer de mama. El modelo alcanzó una precisión del 81.1%, sensibilidad del 95% y especificidad del 98.6%, superando a los patólogos en desempeño. Este enfoque mejoró la precisión diagnóstica y tiene potencial para apoyar a especialistas en la clasificación de células malignas. Similarmente, Aboobacker et al. (2020)) implementaron un modelo U-Net para detectar células malignas y benignas en imágenes de citología de derrame pleural. Este modelo logró una precisión del 96%, recuperación del 96% y especificidad del 97%, con un AUC de 0.97. Su eficiencia lo convierte en una herramienta clave para la detección temprana de cáncer y reducción de falsas alarmas. En (K. Win et al., 2017) se presenta un modelo para mejorar imágenes de núcleos celulares mediante CLAHE y operaciones morfológicas. Evaluado en 35 imágenes de citología pleural, alcanzó un coeficiente de

similitud del 92% y un tiempo promedio de cálculo de 15 minutos por imagen. Este método demuestra su eficacia en diagnósticos automatizados.

La investigación presentada por (S. Wang et al., n.d.), propusieron modelos basados en tomografía computarizada torácica para clasificar derrames pleurales benignos y malignos. Estos modelos integraron segmentación y clasificación, logrando una alta precisión en cohortes de validación, lo que los hace efectivos para el diagnóstico diferencial en oncología clínica. Mientras que en (Li et al., 2021), son integrados aprendizaje profundo y modelos como XGBoost y GBM para predecir derrames pleurales maligno en 726 pacientes. El modelo apilado alcanzó un AUC de 0.991 en cohortes de entrenamiento y validación. Este enfoque optimiza el diagnóstico clínico de cáncer, mejorando la eficiencia predictiva en comparación con métodos tradicionales. En (Xie et al., 2022) se desarrolló un método débilmente supervisado para clasificar células benignas y malignas en imágenes citológicas de pulmón. Este enfoque logró una precisión del 91.67% y un AUC de 0.9526. Su utilidad radica en ayudar a patólogos con distintos niveles de experiencia en el diagnóstico de células cancerosas. El método propuesto será útil y puede ayudar a patólogos con diferentes niveles de experiencia en el diagnóstico de células cancerosas en imágenes en el futuro. (Tosun et al., 2015) proponen un método computacional para analizar núcleos celulares segmentados utilizando cromatina y análisis discriminante. En 34 casos, logró predicciones 100% precisas con validación cruzada, mostrando su eficacia para diagnosticar mesotelioma maligno frente a proliferaciones benignas. Llegaron a la conclusión de que la estructura nuclear de las células mesoteliales por sí sola puede contener suficiente información para separar el mesotelioma maligno de las proliferaciones mesoteliales benignas.

En (Baykal & Doğan, 2017) aplicaron el método Viola-Jones para detectar núcleos celulares en imágenes citológicas. Este enfoque mostró alto rendimiento en la segmentación y clasificación automática de núcleos, facilitando diagnósticos tempranos y precisos en citología serosa. El método fue probado en varias imágenes de citología serosas, los resultados obtenidos muestran que este método tiene un alto rendimiento de detección de núcleos. En (Priyank & Ambika, 2018) exploran soluciones de aprendizaje

profundo para detectar células cancerosas en etapas tempranas. Estas tecnologías mejoran el diagnóstico y la personalización del tratamiento, marcando un avance significativo en oncología. También en (Viktória et al., 2021) realizaron una revisión sistemática sobre técnicas de segmentación y clasificación en citología. Las técnicas de tinción más utilizadas fueron Papanicolaou (130 artículos) y H&E (20 artículos). Este análisis resalta la aplicación de visión por computadora para optimizar diagnósticos citológicos. (Kumar et al., 2022) realizaron una búsqueda de artículos de investigación que emplearon tecnología basada en IA. Enfoques de aprendizaje para la predicción del cáncer. La encuesta deliberó sobre el trabajo realizado por los investigadores y resaltaron las limitaciones de la literatura existente, y realizaron la comparación utilizando varios parámetros como tasa de predicción, precisión, sensibilidad, especificidad, puntuación de datos, tasa de detección, área encubierta, precisión, recuerdo y puntuación F1.

En (Ma et al., 2020) se propuso un algoritmo de segmentación de tres etapas para células tumorales. Estableció un conjunto de datos de células tumorales que consta de 107 conjuntos de imágenes. Este enfoque, aplicado a imágenes teñidas y no teñidas, logró un F1 del 90.40%, destacándose por su precisión en la segmentación y clasificación de regiones celulares. Similarmente, en (Khin et al., 2017) presentaron un algoritmo para la segmentación automatizada de los núcleos celulares en imágenes de citología que tenían células superpuestas. El preprocesamiento para reducir el ruido y mejorar el contraste usando el filtro mediano y Clahe. Los núcleos celulares se segmentan usando K-medias. Se corrigen límites y regiones no nucleas son eliminadas por las operaciones morfológicas. Luego los núcleos de células superpuestas se aíslan por el método de cuenca. Los límites de los núcleos celulares aislados se estiman utilizando el método de ajuste de elipse. El sistema se evalúa en el conjunto de datos local que contiene 35 imágenes de efusiones pleurales de citología con células benignas y cancerosas. En (Jiang et al., 2023), revisaron 120 publicaciones sobre citología computacional basada en aprendizaje profundo. Los enfoques de aprendizaje profundo han demostrado ser efectivos en el análisis de imágenes médicas, estimulando un aumento en las investigaciones relacionadas con la citología. Identificaron avances en clasificación celular y segmentación, destacando desafíos actuales como la falta de datos y proponiendo

direcciones futuras para optimizar diagnósticos citológicos. El estudio resumió los conjuntos de datos públicos, las métricas de evaluación y aplicaciones clave, incluyendo la clasificación celular, la detección de cáncer a nivel de portaobjetos, y la detección y segmentación de núcleos o células. Finalmente, se abordaron los desafíos actuales del campo y se sugirieron direcciones futuras para avanzar en la citología computacional.

(Hooper et al., 2010) precisan que el origen del derrame es un reto clínico, sobre todo porque no hay acceso directo a la cavidad pleural. El proceso de estudio del DP se realiza a través de métodos que incluyen el examen clínico, los estudios radiológicos y tomográficos si los hubiera, y la toracocentesis para obtener el líquido pleural, el que será sometido a procedimientos de citodiagnóstico. Yadanar et al., (2017), sostiene que el examen citológico es indicado para diagnosticar células cancerosas en muestras de derrame pleural porque se considera una herramienta barata, simple, menos invasiva y efectiva. En citología convencional, los anatómicos patólogos examinan todas las células, reconocen las características cito morfológicas anormales y luego presentan su conclusión. La tarea es intensa para el especialista médico, toma tiempo y el trabajo se ve recargado en los países en desarrollo como Perú, debido a la falta de especialistas en los centros de salud. En (Alrafiah, 2022), los modelos de aprendizaje profundo y la inteligencia artificial (IA) logran avances en su capacidad para evaluar e interpretar datos de imágenes, como los de visión por computadora. La primera aplicación de la IA en un laboratorio de patología fue en patología, cuando se creó una prueba de Papanicolaou asistida por computadora. Este fue diseñado para diagnosticar más que para detectar, hubo mucho desacuerdo sobre su amplio uso en muestras clínicas. Sin embargo, las imágenes de portaobjetos completos de histopatología ginecológica y no ginecológica han sido objeto de trabajos recientes de IA.

Rivera et al., (2022), indica que alrededor del 50% de los derrames pleurales (DP) es neoplásico, por lo que es necesario un correcto manejo del DP maligno, se debe tener presente la importancia de detectar células malignas mediante estudio citológico y descartar una patología. En (Ren et al., 2023a), el examen cito patológico es considerado uno de los principales exámenes para el estudio del Derrame Pleural y útil para los pacientes con cáncer avanzado. En muchos casos, el acceso a la muestra de DP es la única

forma para establecer un diagnóstico patológico. En (Naciones Unidas, 2020), consideran que en el sector salud, las tecnologías que utilizan inteligencia artificial ayudan a salvar vidas, diagnosticar enfermedades y prolongar la esperanza de vida. En (Dabeer et al., 2019) menciona que los algoritmos de aprendizaje automático se utilizan para procesar imágenes médicas digitalizadas. La detección manual de una célula cancerosa es una tarea tediosa e involucra error humano, por lo que se aplican mecanismos asistidos por computadora para obtener mejores resultados.

Khin et al., (2017), realizaron la segmentación automatizada de núcleos celulares ya que esas características morfológicas están asociadas a la enfermedad. Para la segmentación usó una clasificación de píxeles que estuvieron basados en una red neuronal artificial (ANN). El tiempo promedio computacional por algoritmo fue de 15 minutos por imagen. Khin et al., (2018), en un estudio comparativo de segmentación mencionan que el 50% de pacientes con cáncer tienen un alto riesgo de desarrollar derrame pleural. Compararon doce métodos y cada uno con tres pasos: preprocesamiento, segmentación y post-procesamiento. Se observó que los tiempos promedios computacionales fueron de 1.08, 36.62, 50.18, 3.30 y 44.03 segundos por imagen. Khin et al., (2018), desarrollaron un método híbrido de selección de características. Aplicaron métodos de ajuste de intensidad y filtrado para mejorar la calidad de la imagen. Utilizaron un clasificador de árboles de decisión para identificar células benignas y malignas. En (Bravo et al., 2021b), implementaron una metodología que clasifica automáticamente el cáncer en leve y agresivo, para esto utilizaron técnicas de procesamiento de imágenes digitales y aprendizaje profundo. Lograron demostrar que el algoritmo diseñado puede usarse como complemento de las herramientas del diagnóstico asistido del cáncer. Mientras que en (Xie et al., 2022), muestran un método de clasificación de células benignas y malignas en imágenes citológicas. Se evaluaron en 60 imágenes de diapositivas completas (WSI). Concluyeron que en un futuro el método puede ayudar a los especialistas que recién inician con diferentes niveles de experiencia en diagnósticos de células cancerosas en líquido pleural.

Lo descrito es sintetizado en la Figura 21, con resumen de las áreas revisadas.

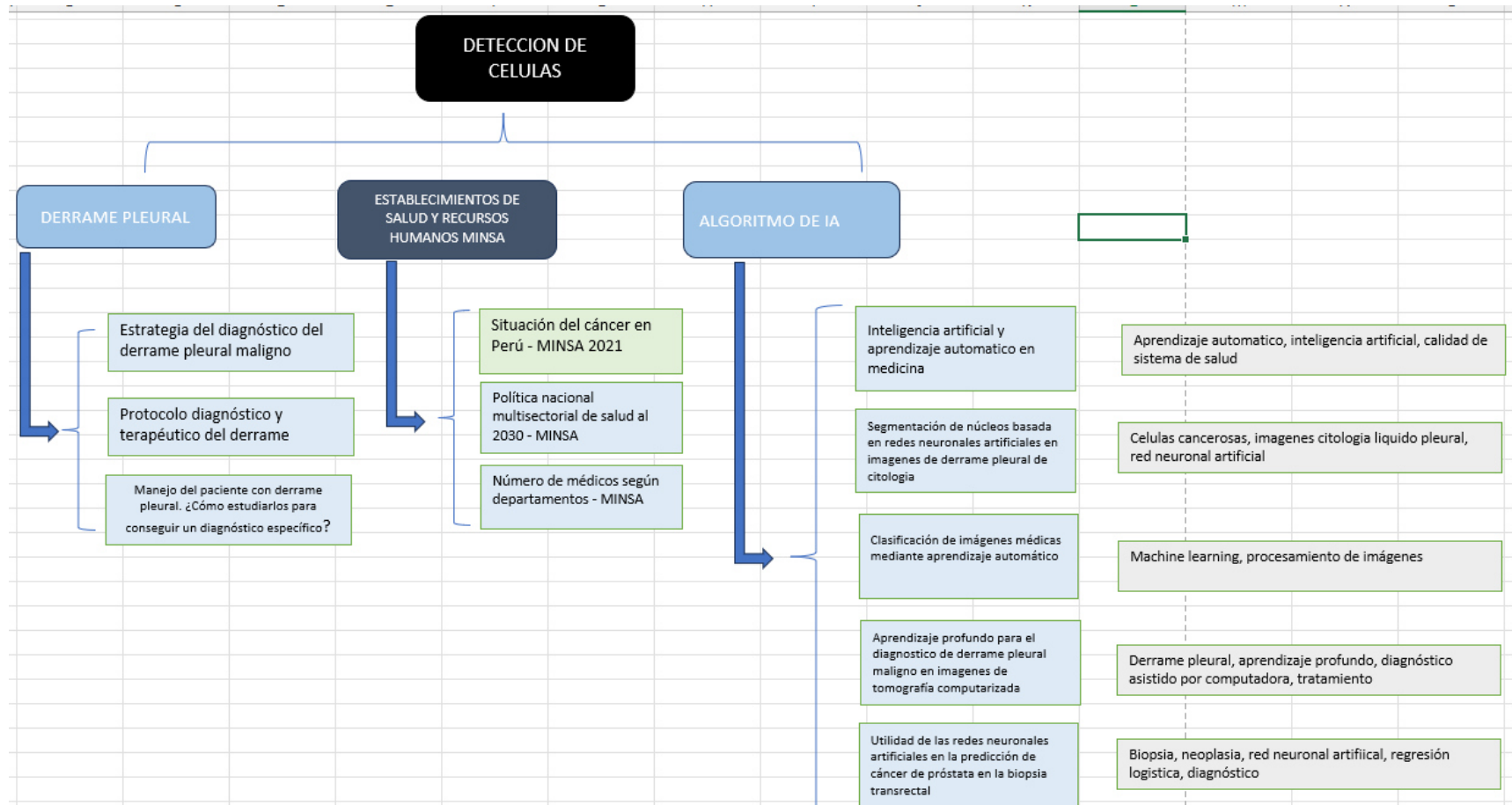


Figura 21. Taxonomía de la revisión de la literatura

3.3. Resultados de la revisión

Se describen estudios que destacan por el uso de técnicas de Machine Learning para la clasificación de células neoplásicas. Se presentan los resultados luego de analizar los artículos, teniendo en cuenta las preguntas de investigación ya definidas, tal y como se visualiza en la Tabla 4.

Tabla 4. Análisis de los hallazgos

Título	Autores	Resumen	Método
K mean clustering based automated segmentation of overlapping cell nuclei in pleural effusion cytology images	Win et al. (2017)	Preprocesan la imagen aplicando ecualización de histograma adaptativa limitada por contraste (CLAHE)	Clasificación CLAHE Red neuronal artificial (ARN) supervisada.
Computer aided diagnosis system for detection of cancer cells on cytological pleural effusion images	Win et al. (2018).	Mejoran la calidad de la imagen, aplicando métodos de ajuste de intensidad y filtrado.	Agrupación lineal Agrupación K-means Red neuronal artificial SA-ANN
A Deep Learning Model for the Automatic Detection of Malignancy in Effusion Cytology	Aboobacker et al. (2020)	DL donde la red aprende directamente a detectar las células malignas en imágenes de citología	U-Net
A machine learning evolutionary algorithm-based formula to assess tumor markers and predict lung cancer in cytologically negative pleural effusions	Elia et al. (2020).	Modelo de aprendizaje automático, ofrecieron una fórmula matemática capaz de expresar la compleja relación existente entre los marcadores expresados en suero y derrame pleural y la presencia de cáncer de pulmón	Técnicas de aprendizaje automático
Deep learning-based segmentation of malignant pleural mesothelioma tumor on computed tomography scans:	Gudmundsson et al. (2020)	Analizaron muestras Entrenaron capas Entrenaron redes profundas con redes inicializadas	ImageNet CNN U-Net

Application to scans demonstrating pleural effusion			
Machine learning applied to near-infrared spectra for clinical pleural effusion classification	Chen et al. (2021)	Desarrollar un método de diagnóstico rápido, conveniente y económico basado en espectroscopia de infrarrojo cercano clínica del derrame pleural	(NIRS) FTIR combinada con Decision Tree SVM
Machine Learning Assisted Real-Time Label-Free SERS Diagnoses of Malignant Pleural Effusion due to Lung Cancer	(Perumal et al., 2022)	Identificar si el cáncer de pulmón de los derrames pleurales se debe a cáncer del pulmón o a otras causas	Clasificador binario PLS_DA Aplicac SERS
Automated Detection, Segmentation, and Classification of Pleural Effusion from Computed Tomography Scans Using Machine Learning	(Sexauer & Shan, 2022)	Entrenaron y evaluaron algoritmos para detectar, segmentar y clasificar derrames pleurales simples y complejos en tomografía computarizada	Red neuronal convolucional profunda U-Net
Deep Learning-Based Classification and Targeted Gene Alteration Prediction from Pleural Effusion Cell Block Whole-Slide Images	(Ren et al., 2023b).	Entrenaron y evaluaron un modelo para diagnóstico de derrame pleural benigno y maligno, para identificar la ubicación primaria del cáncer metastásico común por derrame pleural y la predicción de alteraciones genéticas asociadas con la terapia dirigida	Aprendizaje profundo
Differentiation of malignant from benign pleural effusions based on artificial intelligence	Wang et al. (2022).	Realizaron segmentación y clasificación del derrame pleural benigno y el derrame pleural maligno	Aprend. profundo U-Net 3D U-Net
Automatic deep learning-based pleural effusion segmentation in lung ultrasound images	(Vukovic et al., 2023)	Detectar la presencia de derrame pleural en una ecografía en comparación con un especialista de ecografías experimentado	Modelo de aprendizaje profundo

Deep Learning-Based Computational Cytopathologic Diagnosis of Metastatic Breast Carcinoma in Pleural Fluid	Park et al. (2023)	Evaluaron la precisión diagnóstica de un modelo utilizando portaobjetos citopatológicos, para detectar la citología de derrame pleural maligno asociada con el cáncer de mama	DCNN Inception-Res_Net-V2
Combining Deep Learning And Medical Knowledge to Detect Cardiomegaly and Pleural Effusion in Chest X-rays Diagnosis	(Nguyen et al., 2015)	Desarrollaron modelos de ayuda a radiólogos para identificar en las imágenes de rayos X de tórax: cardiomegalia y derrame pleural.	Algoritmos de segmentación U-Net YOLOv5
Deep learning for diagnosis of malign pleural effusion on computed tomography images	(Ozcelik et al., 2023)	Analizaron imágenes de tomografía computarizada de derrame pleural y examinaron la predicción maligna/benigna comparándolo con los resultados de la citología	Aprendizaje profundo
Development and validation of a machine learning model for differential diagnosis of malignant pleural effusion using routine laboratory data	Wei et al. (2023)	Desarrollaron y validaron un modelo de diagnóstico para la diferenciación temprana entre derrame pleural benigno y maligno basado en datos de laboratorio de rutina.	5 models de ML: XGBoost Regresión logíst Bosque aleatorio SVM MLP
A support vector machine approach for identification of pleural effusion	(Widodo et al., 2024)	Formularon la detección de derrame pleural en imágenes torácicas	Aprendizaje automático: SVM para crear un algoritmo de identificación
Automated detection and segmentation of pleural effusion on ultrasound images using an Attention U-net	(L. Huang et al., 2024)	Modelo de diagnóstico que identifique y segmente automáticamente las áreas de derrame pleural en la ecografía	Attention U-net U-net

Con relación a la PII observamos que existe varios trabajos que utilizan las técnicas de Machine Learning, aplicadas para clasificar células neoplásicas en imágenes citológicas de líquido pleural, abarcan modelos supervisados y enfoques basados en redes neuronales profundas. Algunos estudios han empleado algoritmos clásicos como XGBoost, (Wei et al., 2023) demostrando un alto rendimiento diagnóstico con un área bajo la curva (AUC) de hasta 0,918. En cuanto a redes neuronales profundas, (Park, Chong, Lee, Yim, Seo, Hwang, Kim, Gong, Cho, Yoo, & Choi, 2023) implementaron un modelo de Deep Convolutional Neural Networks (DCNN), obteniendo una precisión superior a la de los patólogos humanos en la clasificación de células neoplásicas. (Chen et al., 2021), integraron la espectroscopia de infrarrojo cercano (NIRS) con algoritmos de aprendizaje automático, logrando resultados óptimos en la diferenciación entre derrames pleurales malignos y benignos. Además, estudios como el de (Elia et al., 2020) emplearon modelos basados en análisis estadísticos avanzados para evaluar métricas como sensibilidad, especificidad y precisión. Técnicas de segmentación y clasificación de imágenes también han sido exploradas, como el uso de coeficientes de similitud de datos (Jaccard y Dice) en redes neuronales convolucionales.

Los estudios muestran que las técnicas más utilizadas en la clasificación de células neoplásicas en imágenes citológicas de líquido pleural incluyen redes neuronales profundas, algoritmos de boosting y métodos de segmentación avanzados, cada uno con diferentes niveles de precisión y aplicaciones clínicas, también se notó que se trabaja frecuentemente con CNNs cuando los datos de entrada son imágenes citológicas de líquido pleural.

En referencia a la PII, las técnicas de Machine Learning aplicadas para diagnosticar células neoplásicas en imágenes citológicas de líquido pleural abarcan modelos supervisados y enfoques basados en redes neuronales profundas. Algunos estudios han empleado algoritmos clásicos como XGBoost, utilizado por (Wei et al., 2023), quien demostró un alto rendimiento diagnóstico con un área bajo la curva (AUC) de hasta 0,918.

En cuanto a redes neuronales profundas, Park et al. (2023) implementaron un modelo de Deep Convolutional Neural Networks (DCNN), obteniendo una precisión superior a la de

los patólogos humanos en la clasificación de células neoplásicas. (Chen et al., 2021), integraron la espectroscopia de infrarrojo cercano (NIRS) con algoritmos de aprendizaje automático, logrando resultados óptimos en la diferenciación entre derrames pleurales malignos y benignos.

Además, estudios como el de (Elia et al., 2020) emplearon modelos basados en análisis estadísticos avanzados para evaluar métricas como sensibilidad, especificidad y precisión. Técnicas de segmentación y clasificación de imágenes también han sido exploradas, como el uso de coeficientes de similitud de datos (Jaccard y Dice) en redes neuronales convolucionales.

En relación a PI2, sobre la disponibilidad de conjuntos de datos de imágenes citológicas para diagnósticos de células neoplásicas, los estudios muestran que las técnicas más utilizadas en la clasificación de células neoplásicas en imágenes citológicas de líquido pleural incluyen redes neuronales profundas, algoritmos de boosting y métodos de segmentación avanzados, cada uno con diferentes niveles de precisión y aplicaciones clínicas, también se notó que se trabaja frecuentemente con CNNs cuando los datos de entrada son imágenes citológicas de líquido pleural.

En la Tabla 5 se realiza una evaluación de los artículos seleccionados según los criterios de clasificación que responden a las preguntas de investigación.

Diversos conjuntos de datos de imágenes citológicas han sido utilizados para el diagnóstico de células cancerosas en estudios recientes. Entre ellos, (Win et al., 2018) emplearon un dataset de 35 imágenes de citología de derrame pleural para evaluar la precisión y recuperación en la clasificación de células neoplásicas. (Baykal & Doğan, 2017) utilizaron un conjunto de 125 imágenes que contenían más de 10.500 células, logrando altos valores de sensibilidad y especificidad. Otro estudio relevante es el de Elia et al., quienes trabajaron con imágenes de 168 pacientes, de los cuales 124 presentaban derrame pleural maligno y 44 no maligno. Park et al. (2023) utilizaron 569 imágenes citológicas de derrame pleural maligno, aplicando modelos de Deep Learning con mejores resultados que el diagnóstico humano. Chen et al. (2021) emplearon un conjunto de 47

muestras de derrame pleural maligno y 35 de benigno, integrando espectroscopia de infrarrojo cercano con aprendizaje automático para mejorar la clasificación. (Aboobacker et al., 2020) analizaron 120 imágenes citológicas, logrando una sensibilidad del 97,15% y una especificidad del 99,80%.

Estos conjuntos de datos han sido fundamentales en la aplicación de técnicas de Machine Learning para mejorar la precisión y fiabilidad en el diagnóstico de células cancerosas en líquido pleural, demostrando la importancia de la recopilación y curación de imágenes citológicas en la investigación médica.

El análisis también observó que los conjuntos de datos para las investigaciones, en su mayoría fueron imágenes de radiografía, imágenes de tomografía e imágenes de ecografía. El conjunto de datos menos representado fue el de imágenes citológicas por lo que se procedió a la construcción de un conjunto de imágenes citológicas de líquido pleural debido a la falencia y ser no accesibles para entrenar los modelos. Asimismo, hace falta estandarizar los conjuntos de datos ya que muchas investigaciones solo trabajan datos locales en cantidad mínima y los datos no son públicos, esto limita la generalización de los resultados.

Tabla 5. Evaluación de los artículos seleccionados

Autor	Dataset	Variables (método de validación)
Win et al. (2017)	35 imágenes de citología de derrame pleural	Rendimiento de 0,95%, 0,86%, 0,90% y 92% en precisión, recuperación, medida F y coeficiente de similitud de datos
Win et al. (2018)	35 imágenes de citología de derrame pleural	Se evalúan cinco métricas de rendimiento
Baykal & Dogan (2017)	125 imágenes de citología de derrame pleural que contenían más de 10.500 células	Sensibilidad del 87,97%, una especificidad del 99,40%, una precisión del 98,70% y una puntuación F del 87,79%

Aboobacker et al. (2020)	120 imágenes de citología	Indice de Jaccard del 90,82 %, una sensibilidad del 97,15 % y una especificidad del 99,80 %
Elia et al. (2020).	Imágenes citológicas de Serum and pleural fluid from 168 patients (124 malignant, 44 non-malignant)	97% de precisión, 98% de sensibilidad, 95% de especificidad, 96% de área bajo la curva, 98% de valor predictivo positivo y 92% de MCC
(Gudmundsson et al., 2020)	154 tomografía computarizada	Mediana de DSC de 0,690
Chen Z. et al. (2021)	47 muestras de DPM 35 muestras de BPE	Precisión 1.0 Kappa 1.0 AUCROC 1.0 Combinar NIRS y aprendizaje automático es un método innovador, rápido y conveniente para la clasificación DP
(Perumal et al., 2022).	Muestras de derrame pleural	Sensibilidad 87% Especificidad 83%.
(Sexauer & Shan, 2022).	160 tomografía computarizada	Entrenamiento 70% Validación 30%
(Ren et al., 2023a)	1.321 casos de derrame pleural	Identificación de derrames pleurales benignos y malignos (0,932 AUC ROC) y la ubicación primaria del cáncer metastásico (0,910 AUC).
Wang et al. (2022)	918 imágenes de tomografía computarizada	Coeficiente de DSC 87.6% Coeficiente de Jaccard 82.2% Precisión 99% Sensibilidad 83% Distancia de Hausdorff 95%
(Vukovic et al., 2023).	3041 imágenes de ecografías	Coeficiente DSC 0,70 entre el algoritmo y los expertos. Coeficiente DSC 0,61 entre los expertos
Park et al. (2023)	569 imágenes citológicas de	El modelo DCNN obtuvo mayor precisión, sensibilidad y especificidad en

	derrame pleural maligno	comparación con los patólogos (81,1% frente a 68,7%, 95,0% frente a 72,5% y 98,6% frente a 88,9%, respectivamente).
(Nguyen et al., 2015)	Imágenes de radiografía de tórax	Precisión/recuperación del 78 %/93 % para cardiomegalia y del 72 %/93 % para derrame pleural
(Ozcelik et al., 2023).	408 imágenes de tomografía	(PPD: 93,3%, NPD: 86,67%, Sensibilidad: 87,5%, Especificidad: 92,86%)
Wei et al. (2023)	2352 pacientes con derrame pleural, incluida la edad, el sexo y los parámetros de laboratorio	XGBoost mostró el mejor rendimiento diagnóstico para MPE (área bajo la curva (AUC): 0,903, 0,918 y 0,886 en las cohortes de entrenamiento, validación y prueba, respectivamente
(Widodo et al., 2024).	100 imágenes de radiografía de tórax	Precisión 96%
(L. Huang et al., 2024)	848 imágenes de radiografía de tórax	Attention U-net tuvo un mejor desempeño en la identificación de imágenes negativas con 91% que U-net con 86%. Además, el fué más preciso en detección de la región de derrame pleural.

3.4. Comentarios de la revisión

A partir de cada artículo de investigación se obtuvo información referente a: número de imágenes, división del conjunto de datos para entrenamiento, validación y prueba, las técnicas de preprocesamiento, métricas como accuracy, precisión, recall, F1, AUC, validación utilizada y conclusiones.

Se definieron dos preguntas de investigación para desarrollar la revisión de la literatura y colaborar con los investigadores a encontrar enfoques y tendencias actuales en este campo. Respecto a la P1 los modelos utilizados en las investigaciones referente al uso de técnicas usadas de Machine Learning se pudo observar que destacan en su uso los modelos U_Net (Convolutional Networks for Biomedical Image Segmentation), Support vector machine (SVM) y el modelo CNN (Convolutional Neural Network). Para la PI2 sobre el conjunto

de datos utilizado en las investigaciones, se verificó que el tipo de datos más usado son las imágenes de citologías, tomografías y radiografía. Los hallazgos favorecen futuras investigaciones en modelos y técnicas de Machine Learning, en apoyo a los profesionales de la salud y reduce el tiempo de entrega de resultados. Además, permite el uso de tecnologías para el apoyo diagnóstico y contribuir a un tratamiento oportuno en pacientes diagnosticados con derrame pleural.

El análisis muestra las siguientes tendencias: a) conjunto de datos pequeños y desbalanceados, b) carencia de conjunto de datos y c) escasa validación externa. Estas limitaciones justifican la adopción de transfer learning, data augmentación en tres niveles y construcción del conjunto de datos en la presente investigación.

Se puede indicar que los trabajos revisados ofrecieron una visión panorámica de los datos, sin embargo, quedan abiertas las posibilidades de temas que no han sido tratados en ese campo como la aplicación de modelos de Transfer Learning, que se aplican en esta investigación.

CAPÍTULO 4. METODOLOGÍA DE LA INVESTIGACIÓN

La metodología combina el marco de Design Science Research (DSR) con un pipeline técnico de ML. Se describen las arquitecturas seleccionadas, estrategias de fine-tuning, esquema de validación, métricas utilizadas y pruebas aplicadas.

Se describe la forma que se trabajó la investigación, cuyo objetivo principal fue diseñar un modelo para la clasificación de células neoplásicas en imágenes de citología. Se realizaron evaluaciones en todos los procesos asegurando que cumplan con la precisión y tiempo oportuno de proceso, para lograrlo se construyó el artefacto (modelo) para luego ser entrenado y probado con el conjunto de datos.

La metodología se estructura en dos componentes principales: la metodología de investigación general para el desarrollo de la tesis y la metodología específica para el diseño del modelo, basada en los principios de la ciencia del diseño y complementando con los principios metodológicos de (Montgomery, 2004).

Según (Hevner et al., 2004), la Ciencia del Diseño es un paradigma de investigación que busca crear y evaluar artefactos tecnológicos diseñados para resolver problemas organizacionales identificados. Este enfoque se complementa con la metodología de investigación mixta propuesta por (Hernández-Sampieri et al., 2014), quien establece que "los métodos mixtos representan un conjunto de procesos sistemáticos, empíricos y críticos de investigación e implican la recolección y el análisis de datos cuantitativos y cualitativos" (p. 534).

Paradigma de Investigación

El paradigma adoptado es el pragmatismo, el cual, según Creswell (2014), permite la utilización de múltiples métodos de investigación para abordar un problema específico. En el contexto de la Ciencia del Diseño, este paradigma es particularmente relevante ya que, como señalan Peffers et al. (2007), "la investigación en ciencia del diseño es inherentemente un proceso de resolución de problemas" (p. 45).

4.1. Metodología general de la investigación

Los hallazgos de la RSL, desarrollado en el Capítulo 3, orientan a la selección de arquitecturas y modelos experimentales descritos en esta parte de la investigación.

Tipo y diseño de la investigación

La investigación es de tipo aplicada, porque busca aplicar el conocimiento tecnológico para el desarrollo y solución de un problema de la vida real, el cual se busca resolver mediante la utilización de algoritmos que servirán para mejorar la toma de decisiones de los expertos en el campo de salud.

De acuerdo con Montgomery (2017), la investigación puede clasificarse en dos categorías metodológicas: estudios observacionales y estudios experimentales. En los estudios observacionales, el investigador observa y registra información sin interferir en el proceso natural, mientras que, en los estudios experimentales, el investigador manipula deliberadamente una o más variables para observar su efecto sobre una variable de respuesta.

La presente investigación se clasifica como:

- Aplicada: Porque está orientada a resolver problemas prácticos en el diagnóstico de células neoplásicas.
- Exploratoria-descriptiva: Porque explora tres arquitecturas algorítmicas al describir su comportamiento
- Experimental: Porque coincidiendo con la definición de (Montgomery, 2004), constituye un estudio experimental verdadero donde se manipulan controladamente las variables algorítmicas (hiperparámetros, arquitecturas, técnicas de preprocesamiento) para evaluar su efecto sobre las variables de respuesta (precisión, recall, F1-score). Además, los experimentos permiten establecer relaciones causales y son superiores a los estudios observacionales para

validar la efectividad de intervenciones o tratamientos, en este caso, algoritmos de clasificación.

En el marco de la Ciencia del Diseño, (March & Smith, 1995) establecen que este tipo de investigación "está orientada hacia la creación de artefactos para resolver problemas prácticos" (p. 253), lo que coincide con el carácter aplicado de la presente investigación.

Unidad de análisis

Está compuesta por las 1292 imágenes del conjunto de datos contruidos exclusivamente para la presente investigación. Corresponde a un trabajo multidisciplinario con el HNCH. Asimismo, un segundo conjunto de datos con 693 imágenes para la evaluación y generalización del modelo.

Población de estudio

Conformada por el conjunto de datos propios construido con imágenes del HNCH que comprende imágenes de citología de líquido pleural y un conjunto de datos adicional obtenido de un repositorio para evaluar la generalización del modelo.

Tamaño de muestra

Conjunto de Datos del HNCH: 1292 imágenes

Conjunto de Datos Repositorio público: 693 imágenes

Selección de muestra

Las muestras del conjunto de datos propios fueron distribuidas en 70% entrenamiento y 30% evaluación. Las muestras del segundo conjunto de datos fueron asignadas al 100% para evaluación del modelo.

Técnicas de recolección de datos

Se ha realizado el procedimiento de búsqueda de la información del tema relacionado en el estado del arte; se realizó una búsqueda directa en el HNCH donde luego de una serie

de procedimientos administrativos se obtuvo la autorización del Comité de Ética del HNCH para coleccionar imágenes y construir el conjunto de datos propios exclusivo para la presente investigación, específicamente para entrenamiento y evaluación de los experimentos del modelo.

Análisis e interpretación de la información

Se realizaron las siguientes actividades para el análisis e interpretación de resultados:

- Construcción de los 3 modelos
- Entrenamiento de cada modelo
- Entrenamiento de las diferentes combinaciones de modelos
- Análisis de las métricas obtenidas de todos los entrenamientos
- Validación del modelo con el segundo conjunto de datos
- Análisis y determinación del mejor modelo

4.2. Diseño del Artefacto: Ciencia del Diseño

Para el desarrollo del algoritmo de clasificación se adopta la metodología de Design Science Research (DSR), propuesta por Hevner et al. (2004) y refinada por Peffers et al. (2007). Esta metodología es particularmente apropiada para el desarrollo de artefactos tecnológicos en el contexto de sistemas de información y ciencias de la computación. Este paradigma de investigación se orienta hacia la creación de artefactos innovadores con el propósito de solucionar problemas y contribuir al desarrollo del conocimiento científico. Los artefactos son creaciones humanas diseñadas para realizar tareas específicas o cumplir funciones determinadas. La ciencia del diseño se centra en el estudio de estos artefactos con el fin de optimizar soluciones existentes o desarrollar nuevas alternativas a problemas aún no resueltos (Hevner et al., 2004).

En el ámbito de la ciencia del diseño como disciplina de investigación, el eje se encuentra en la creación o modificación de artefactos para mejorar la eficiencia y efectividad de las soluciones actuales (March & Smith, 1995).

Esta área ha tomado relevancia como una respuesta a la orientación conductual predominante en la investigación de sistemas de información, buscando complementar esta perspectiva con un enfoque más práctico y orientado al diseño. En (Hevner & Chatterjee, 2010) se definen la investigación en la ciencia del diseño como un paradigma de investigación cuyo principio fundamental de la investigación en ciencias del diseño es que el conocimiento y la comprensión de un problema de diseño y su solución se adquieren en la construcción, aplicación y análisis de un artefacto.

En esta investigación, el artefacto es un modelo basado en aprendizaje profundo diseñado para clasificar células neoplásicas y se encontró que la DSR establece principios metodológicos sólidos que garantizan un desarrollo y una evaluación de esta técnica, asegurando no solo su efectividad en la práctica, sino también su aporte al conocimiento científico en el ámbito del diagnóstico médico.

En comparación con las ciencias del comportamiento, que buscan descubrir verdades mediante la verificación de hipótesis, la DSR se centra en la solución de problemas a través de la creación y evaluación de artefactos (Hevner & Chatterjee, 2010), no obstante, es fundamental destacar que el uso de DSR no implica abandonar los principios fundamentales de la investigación científica, sino que busca establecer una sinergia que permita desarrollar un artefacto mientras se alcanzan los objetivos investigativos y se comprueban las hipótesis formuladas. Estas últimas, en lugar de verificarse a través de enfoques hermenéuticos o pruebas estadísticas típicas de los paradigmas de investigación convencionales, serán comprobadas a partir de los resultados obtenidos al aplicar una metodología específica de DSR.

El concepto de artefacto es central en esta disciplina, diferenciándose de otros métodos de investigación cualitativos o cuantitativos. Un artefacto es cualquier objeto construido por el ser humano, a diferencia de aquellos que ocurren de manera natural (Simon, 1996). La

obra *The Science of the Artificial* de Simon (1996) establece la base conceptual de este enfoque, describiendo los artefactos como fenómenos empíricos de naturaleza artificial en contraposición a los fenómenos naturales.

Los métodos de la ciencia del diseño han captado la atención de la comunidad de sistemas de información, especialmente tras el artículo publicado por Hevner et al. (2004). Este paradigma de investigación propone que el diseñador responde a problemas humanos mediante la creación de artefactos innovadores, generando así nuevos conocimientos que contribuyen al desarrollo de la evidencia científica. Los artefactos diseñados no solo son útiles, sino también esenciales para comprender mejor los problemas que buscan resolver.

Estos artefactos transcurren por dos actividades básicas de la investigación en la DSR: la construcción y la evaluación. Cada actividad posee subactividades para garantizar la calidad de la investigación.

Por esta razón (Peppers et al., 2007), proponen la Metodología de Investigación en Ciencias del Diseño (DSRM), un marco para desarrollar y presentar trabajos de investigación en ciencias del diseño, en especial en el área de los sistemas de información. Las cuales están compuestas por seis actividades que orientar el proceso de investigación como se muestra en la Figura 22.

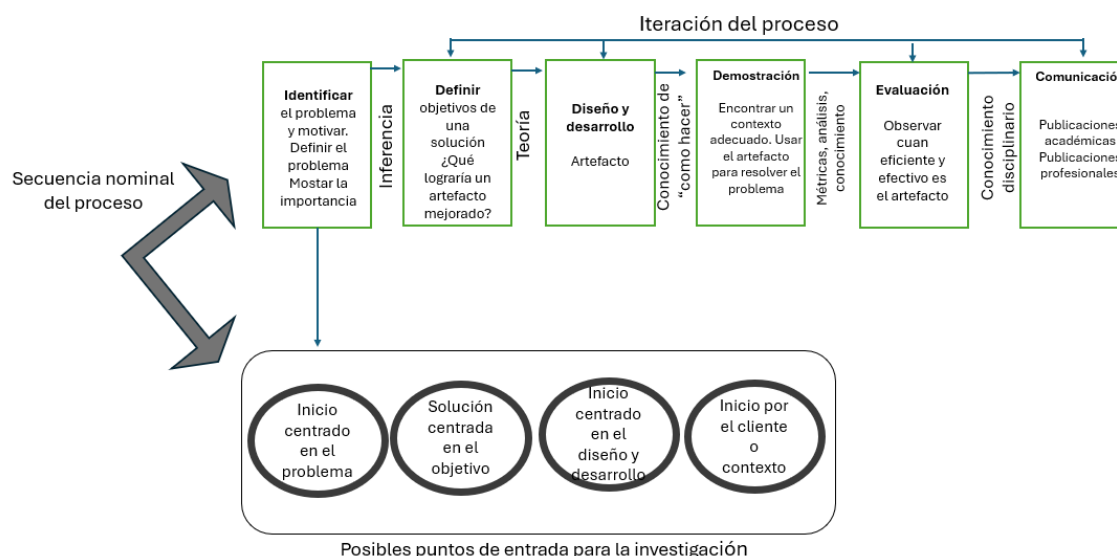


Figura 22. Metodología de Ciencia del Diseño DSR

Fuente. Adaptado de (Peffers et al., 2007).

4.2.1. Identificación del problema y sus motivaciones

Al contrario que las ciencias del comportamiento, que demuestran hipótesis para descubrir verdades, la DSR se centra en resolver problemas mediante la construcción y evaluación de artefactos. No busca abandonar los principios científicos, sino integrarlos para desarrollar artefactos mientras se alcanzan los objetivos de investigación. En este enfoque, las hipótesis no se verifican con métodos tradicionales, sino a través de la aplicación de una metodología específica para DSR (Peffers et al., 2007).

a) Objetivos

En la investigación DSR, los objetivos difieren de los de las ciencias naturales, ya que no buscan únicamente generar conocimiento teórico, sino desarrollar soluciones prácticas para problemas reales. Según (Pries-Heje & Baskerville, 2008), los objetivos en DSR se centran en la funcionalidad y los meta-requisitos del sistema. Peffers et al. destacan que estos pueden ser cuantitativos, mejorando métricas existentes, o cualitativos, abordando problemas aún no tratados. La formulación de estos objetivos requiere un análisis del estado del problema y la efectividad de soluciones previas. En este contexto, la DSR se

orienta hacia la producción de conocimiento aplicable, con un enfoque pragmático que contribuye activamente a la resolución de problema.

b) Diseño y desarrollo

Se considera la construcción del artefacto, pueden ser constructos, modelos, métodos o instantaneos. En esta parte se determina la funcionalidad, arquitectura y creación del artefacto, Se requiere conocer la teoría que sustenta la solución.

c) Demostración

Se implementan los experimentos, simulaciones, pruebas y demás actividades para que se demuestren con eficacia la resolución de los problemas para lo que construyó el artefacto.

d) Evaluación

Consiste en analizar su desempeño como solución al problema planteado, comparando los resultados obtenidos con los objetivos establecidos. Este proceso requiere la aplicación de técnicas de análisis y métricas relevantes. Según los hallazgos, los investigadores pueden decidir si es necesario iterar en la fase de diseño y desarrollo para mejorar la solución o proceder a la fase de comunicación, documentando las mejoras y proponiendo direcciones para futuras investigaciones.

e) Comunicación

Se difunden los resultados de la investigación, destacando la naturaleza del problema, la utilidad y novedad del artefacto, así como el rigor y la eficacia de su diseño. Esta difusión debe adaptarse tanto a la comunidad académica, mediante formatos tradicionales de investigación, como a profesionales del sector. Además, la comunicación requiere comprender la cultura disciplinaria, la cual influye en la producción, transmisión e intercambio del conocimiento dentro de una comunidad académica.

4.3. Justificación de la metodología DSR

Se adopta Design Science Research (Hevner et al., 2004; Peffers et al., 2007) por su orientación a la construcción y evaluación de artefactos tecnológicos.

La Metodología de Investigación en Ciencias del Diseño (DSRM) es utilizada comúnmente para resolver problemas diversos. En los últimos años, se ha visto la modernización del sistema de salud, también se ha vuelto más complejo. Relacionando diferentes partes interesadas, inversión financiera, estándares clínicos, protocolos de salud y regulaciones gubernamentales. Esta complejidad se vuelve aún mayor a medida que cada paciente, cada enfermedad y condiciones subyacentes vienen con sus propias características que influyen directa o indirectamente entre sí. Tal es así, que se encontraron numerosos trabajos que usaron DSRM. A continuación, se resume las investigaciones realizadas en los últimos tres años de temas de salud relacionados con imágenes médicas que aplican la DSRM.

(X. Wang et al., 2025) propone un artefacto de detección de lesiones en imágenes médicas, mencionando explícitamente la metodología DSR para soporte de decisión, usando imágenes médicas para detectar lesiones.

(Pumplun et al., 2023), aplicó la metodología DSR para desarrollar un sistema de soporte a decisiones clínicas basado en IA explicable en el diagnóstico de nódulos pulmonares. Siguiendo ciclos iterativos de diseño, construcción y evaluación con médicos, el artefacto integró principios de usabilidad, transparencia y confiabilidad, fortaleciendo la adopción clínica de herramientas basadas en aprendizaje profundo.

(Berkhout et al., 2024), mediante la metodología DSR, diseñó y evaluó un sistema para descubrir patrones de decisión en datos de soporte clínico. Los ciclos iterativos de diseño y evaluación permitieron optimizar procesos hospitalarios, integrando minería de datos y retroalimentación médica para mejorar la eficiencia, precisión y personalización en la toma de decisiones clínicas.

(Gottumukkala, 2023), desarrolló una solución de intercambio automatizado de datos clínicos usando la metodología DSR, asegurando interoperabilidad mediante estándares HL7/FHIR. El artefacto fue diseñado, demostrado y evaluado en entornos reales, mostrando mejoras significativas en precisión y eficiencia, y validando la aplicabilidad del DSR en salud digital y sistemas inteligentes.

(Mdletshe et al., 2023), aplicó la metodología DSR para crear un artefacto educativo en radiología, diseñado y evaluado en ciclos iterativos con estudiantes de ciencias médicas. La investigación integró innovación tecnológica y necesidades pedagógicas, demostrando cómo el DSR garantiza efectividad, usabilidad y transferencia de conocimiento en la enseñanza de imágenes médicas.

En esta investigación, el artefacto es un modelo de aprendizaje profundo para citología pleural. DSR permite articular relevancia (problema real), rigor (estado del arte) y evaluación (métricas de desempeño), en línea con los principios de investigación doctoral. Esta elección no sustituye al pipeline técnico de ML, sino que lo enmarca. El estado del arte justifica las técnicas empleadas: aprendizaje por transferencia con ResNet50V2, DenseNet121, InceptionV3; augmentación; fine-tuning; y ensembles. La combinación de DSR y el pipeline experimental garantiza tanto contribución científica como validez aplicada.

CAPÍTULO 5. CONSTRUCCIÓN DEL CONJUNTO DE DATOS DE IMÁGENES CITOLÓGICAS DE LÍQUIDO PLEURAL

La calidad de los datos es fundamental para el entrenamiento de los modelos por lo tanto deben etiquetarse con precisión e incluir diversidad morfológica. La creación de un conjunto de datos fue una de las primeras acciones de esta investigación (López-Córdova et al., 2025) y tratándose de imágenes citológicas de líquido pleural, fue una tarea necesaria y relevante al no encontrar conjuntos de datos de acceso libre para entrenar los modelos de Inteligencia Artificial. Para este fin, es necesario seguir una metodología estructurada que asegure la calidad, representatividad y diversidad de las imágenes (Phoulady & Mouton, 2018).

En esta parte se describe la construcción de un conjunto de datos de imágenes citológicas de líquido pleural para ser utilizado en el entrenamiento de modelos de inteligencia artificial (IA). Se construyó un conjunto de imágenes citológicas validadas por dos patólogos y clasificadas en cinco categorías diagnósticas según el Sistema Internacional (TIS) para informar líquidos serosos. Se describe el proceso de construcción completo, desde la captura, etiquetado y preprocesamiento de las imágenes. Se aplicaron técnicas de mitigación de ruido como normalización de tinción, filtro mediano 3×3 , operaciones morfológicas, umbral de Otsu para remover fondo, y descarte de imágenes desenfocadas. Estas medidas aseguran calidad diagnóstica y consistencia del nuevo conjunto de datos que luego fue utilizado en el entrenamiento de los modelos de esta investigación. Este conjunto de datos pretende ser de apoyo para las investigaciones y utilización en el desarrollo de modelos de aprendizaje profundo para mejorar los diagnósticos.

5.1. Construcción del conjunto de datos de imágenes citológicas de líquido pleural.

El conjunto de datos que da soporte a esta investigación ha sido creado a partir de estudios citológicos con coloración Papanicolau y de aplicación para metástasis de tumores sólidos de tipo epitelial (carcinomas). Cada estudio tiene su diagnóstico cito patológico realizado por el equipo de médicos patólogos. Las muestras pertenecen al Servicio de Anatomía Patológica del Hospital Nacional Cayetano Heredia (Lima-Perú) y forman parte del archivo de estudios citológicos de dicho servicio que son utilizados con fines de docencia y de investigación. El protocolo de investigación fue aceptado por el Departamento de Patología Clínica y Anatomía Patológica y aprobado por el Comité Institucional de Ética en Investigación del Hospital Nacional Cayetano Heredia el 12 de diciembre de 2024, con el código 102-2024 (Ver anexo 1, anexo 2 y anexo 3).

El microscopio

Un microscopio es el instrumento que permite observar objetos que son demasiado pequeños como para ser vistos por la vista del hombre. El término *microscopio* forma parte de dos conceptos, “*micro*” que es equivalente a “*pequeño*” y “*scopio*” que significa “*observar*”, por lo que se entiende como una observación pequeña o en menor grado. El microscopio es un instrumento óptico que aumenta la capacidad de observación a niveles de acercamiento tal que hasta hace posible el análisis de partículas. La imagen que produce ayuda a la investigación de la composición de los objetos, por tanto, estudiar y analizar los objetos pequeños se denomina “microscopia”. Como es común en la mayoría de las investigaciones el campo de visión de un microscopio óptico se aumenta a más de 60 (Aubreville et al., 2020)

La microscopía es la observación de objetos muy pequeños bajo grandes aumentos. Los aparatos que se usan para ello se denominan microscopios. En la medicina se usan la microscopía especialmente para analizar tejidos, células, componentes sanguíneos y microorganismos. Las imágenes citológicas se analizan utilizando un microscopio óptico con cámara digital integrada, marca Olympus, equipado con un ocular de 10x y un

objetivo de 60x, lo que da como resultado un aumento de 600x. Las imágenes se capturan a diferentes aumentos (10x, 60x) permitiendo el análisis detallado de las células. El alto nivel de aumento garantiza que los detalles celulares relevantes para el diagnóstico sean claramente visibles. Para guardar las imágenes sin perder la calidad, el formato usado fue PNG porque preserva la calidad visual sin compresión excesiva.

El microscopio óptico, tal como ilustra la Figura 23, capturará todas las imágenes posee una cámara incorporada y un software que se encuentra instalado en la computadora del servicio de anatomía patología, el cual hace posible la toma y almacenamiento de la imagen.



Figura 23. Microscopio óptico del HNCH

Fuente: Fotografía tomada por la investigadora.

El conjunto de datos construido está formado por 2640 imágenes de alta calidad en formato PNG, cada una con una resolución de 4140 x 3096 píxeles que representa la

resolución de la imagen, el tamaño de las imágenes varía entre 11 a 12 MB. Para construir el conjunto de datos se realizaron las siguientes etapas:

5.2. Captura de imágenes

Esta tarea fue realizada por los especialistas, se coloca la lámina en el microscopio con un lente objetivo de 20 de aumento, se ubica el área con células sospechosas para la categoría diagnóstica que se está buscando. Luego se cambia el objetivo a 60 de aumento y se enfoca la mejor imagen a capturar con el software propio del microscopio instalado en la computadora del servicio de anatomía patológica. El microscopio posee una cámara que se acciona a través de la computadora (la cámara tiene un cable que se conecta directo con la computadora). Con el objetivo en 60 de aumento se señala en el campo visual del microscopio las células motivo de la búsqueda. Una vez fijada la imagen en el campo visual a capturar, se visualiza la imagen a través de la computadora para ser capturada. El enfoque final se procesa en la pantalla de la computadora, se realiza la captura de la imagen a color y se almacena en la carpeta correspondiente a la categoría diagnóstica.

5.3. Categoría diagnóstica de las imágenes

Las imágenes fueron revisadas por dos médicos patólogos para verificar que hayan sido etiquetadas en sus respectivas clases. Los dos especialistas patólogos con experiencia de 20 años y 8 años respectivamente. Las imágenes están organizadas en etiquetas con su respectiva categoría diagnóstica según la presencia o ausencia de malignidad. Médicos patólogos expertos en citología identificaron las imágenes, etiquetando diferentes tipos de células encontradas de acuerdo con el TIS (AUS células atípicas, MAL células malignas, ND no diagnóstico, NFM células negativas para malignidad y SFM células sospechosas para malignidad) tal como se visualiza en la Tabla 6. Estas etiquetas son esenciales para el entrenamiento de los modelos computacionales.

Tabla 6. Categorías diagnósticas de acuerdo con el TIS.

SISTEMA INTERNACIONAL PARA INFORMAR CITOLOGIA DE LIQUIDOS SEROSOS: RIESGO DE MALIGNIDAD	
ATIPIA DE SIGNIFICADO INDETERMINADO	AUS
MALIGNO	MAL
NEGATIVA PARA MALIGNIDAD	NFM
NO DIAGNOSTICO	ND
SOSPECHOSAS DE MALIGNIDAD	SFM

5.4. Preprocesamiento de las imágenes

Se inició con la implementación de un programa en código Python para ejecutar las diferentes tareas del procesamiento de imágenes como: 1) nombrar y enumerar las imágenes en sus respectivas carpetas, 2) configurar la ruta de las cinco clases, 3) redimensionar las imágenes a un tamaño de 250x250 pixeles para facilitar el entrenamiento y 4) convertir a escalas de grises.

5.4.1. Etiquetado de imágenes para organizar el conjunto de datos

Se elaboró un código en Python para organizar los archivos, etiquetar imágenes con nombres descriptivos para identificación y clasificación de acuerdo con su categoría diagnóstica. En la Figura 24 podemos observar el código para organizar los archivos y en la Figura 25 las imágenes renombradas y numeradas.

```

import os
from PIL import Image

# Ruta donde se encuentran las imágenes
ruta = "D:/DataSet Pleural/SFM"

# Obtener la lista de archivos en la ruta especificada
archivos = os.listdir(ruta)

# Filtrar solo los archivos de imagen
imagenes = [archivo for archivo in archivos if archivo.lower().endswith(('.png', '.jpg', '.jpeg', '.gif', '.bmp'))]

# Ordenar las imágenes alfabéticamente
imagenes.sort()

# Contador para mantener el número de imagen
contador = 1

# Iterar sobre cada imagen
for imagen in imagenes:
    # Generar el nuevo nombre de imagen con el formato especificado
    nuevo_nombre = f"SFM_{contador:03d}.png" # Puedes cambiar la extensión si es necesario

    # Evitar sobrescribir imágenes existentes
    while nuevo_nombre in archivos:
        contador += 1
        nuevo_nombre = f"SFM_{contador:03d}.png"

    # Renombrar la imagen
    os.rename(os.path.join(ruta, imagen), os.path.join(ruta, nuevo_nombre))

    # Incrementar el contador para la siguiente imagen
    contador += 1

print("Imágenes renombradas con éxito.")

```

Figura 24. Código para organizar archivos

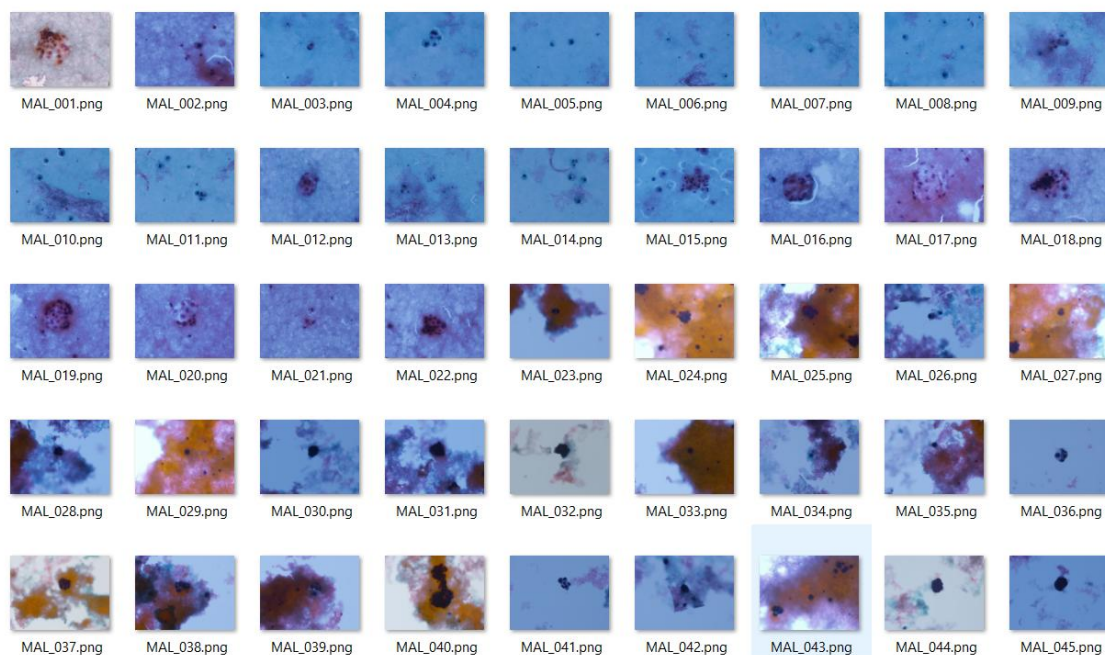


Figura 25. Imágenes renombradas y numeradas

5.4.2. Dimensionamiento de las imágenes

La adquisición de imágenes es un proceso de captura, consiste en convertir un objeto o una escena, en una representación apta para ser procesada por una computadora. Tal adquisición o conversión puede ser realizada por medio de un escáner o cámara fotográfica. El resultado varía dependiendo del proceso de digitalización usado y del método de codificación (Martínez et al., 2016).

Se utilizaron diferentes técnicas que aplican una operación matemática sobre la imagen. En este caso nuestras variables de entrada son los valores de los píxeles. El objetivo de este proceso es modificar las imágenes para mejorar su representación en un espacio de características y facilitar el proceso de aprendizaje (Martínez et al., 2001)

Para estandarizar a un tamaño fijo las imágenes ya que a veces se encuentran en diferentes tamaños, las imágenes del conjunto de datos se redimensionaron a 250 x 250 píxeles, utilizando la función RESIZE de Python.

5.4.3. Conversión a blanco y gris

Luego se procedió a la conversión a blanco y grises, como las imágenes fueron obtenidas con diferentes tinciones, el color de fondo de la imagen era variado y no ayudaba al aprendizaje del modelo, por tanto, se decidió pasar las imágenes a escalas de grises con la función IMREAD_GRAYSCALE de Python. En la Figura 26 se puede apreciar dos imágenes de las categorías MAL (malignas) y NFM (benignas) en escalas de grises.

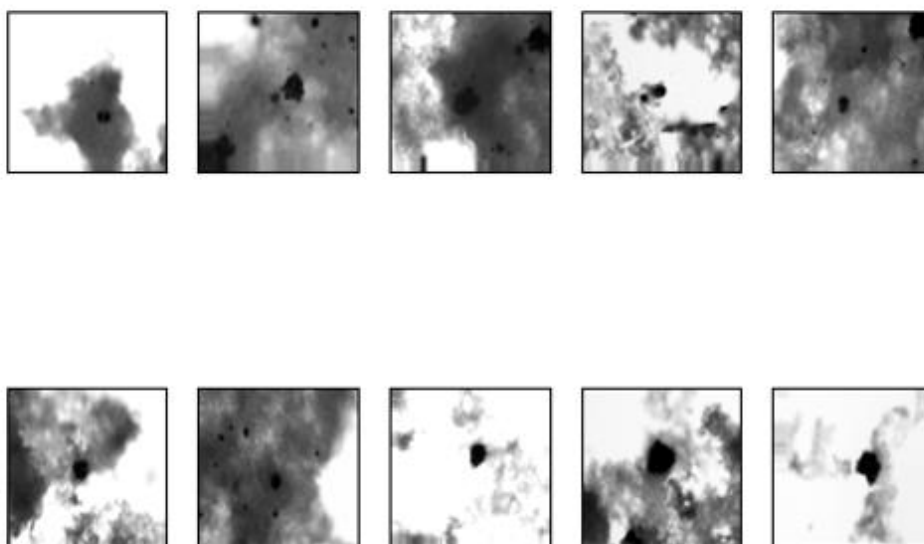
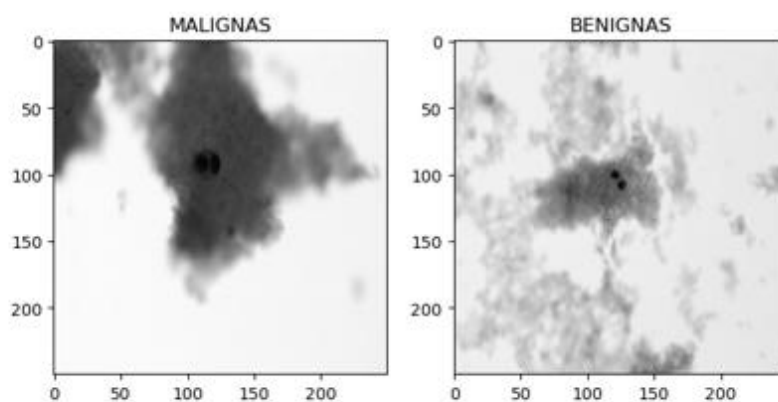


Figura 26. Imágenes del conjunto de datos en escala de grises

Al culminar el preprocesamiento tenemos el conjunto de datos tal como lo muestra la Figura 28 de imágenes clasificadas por categoría diagnóstica.

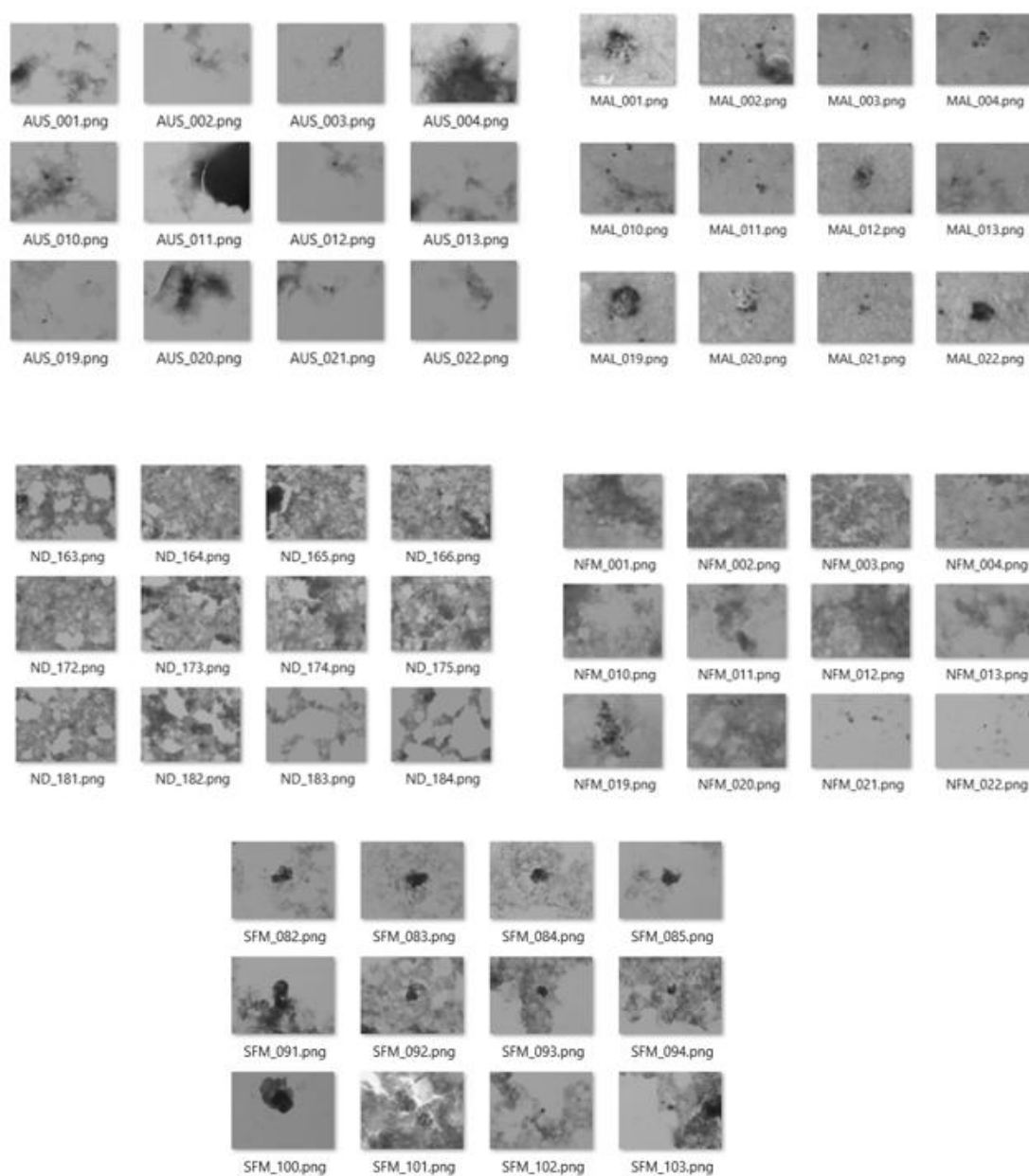


Figura 27. Imágenes nombradas y numeradas por categoría diagnóstica

En la Tabla 7 se muestra el número de imágenes por categoría diagnóstica que fueron obtenidas en el periodo de la investigación. Se puede observar que las categorías MAL y NFM tenían cantidades similares (estaban balanceadas entre ellas), sin embargo, las categorías AUS, ND y SFM tenían cantidad menores y diferentes entre ellas.

Tabla 7. Imágenes nombradas y numeradas por categoría diagnóstica

SISTEMA INTERNACIONAL PARA INFORMAR CITOLOGIA DE LIQUIDOS SEROSOS: RIESGO DE MALIGNIDAD	Categoría	Total de imágenes
ATIPIA DE SIGNIFICADO INDETERMINADO	AUS	358
MALIGNO	MAL	645
NEGATIVA PARA MALIGNIDAD	NFM	647
NO DIAGNOSTICO	ND	463
SOSPECHOSAS DE MALIGNIDAD	SFM	527
TOTAL		2640

En la actividad médica, cuando se diagnostica la categoría MAL o SFM, es necesario indicar otros exámenes adicionales para confirmación de diagnóstico.

En la Figura 28 se muestra imágenes del conjunto de datos construido de las muestras contenidas en los portaobjetos completos (WSI) del servicio de anatomía patológica del Hospital Nacional Cayetano Heredia (Lima Perú) por categoría diagnóstica. Siendo: AUS-Atipia de significado indeterminado, ND-No diagnóstico, NFM-Negativa para malignidad, SFM-Sospechosa para malignidad y MAL-Maligna.

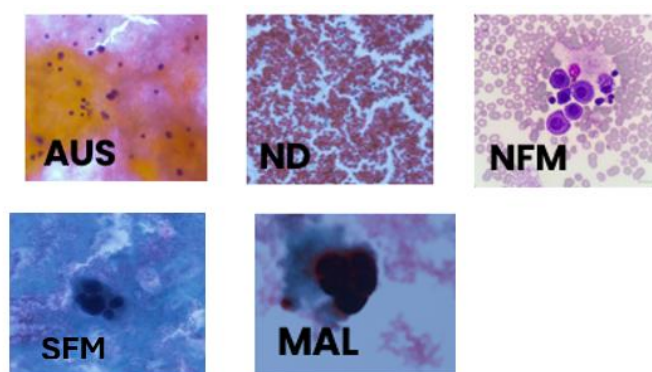


Figura 28. Imágenes del conjunto de datos construido.

Propósito de la construcción

El conjunto de datos fue construido como parte de esta investigación al tomar conocimiento que en el Perú no se encontró un conjunto de datos de imágenes citológicas de líquido pleural. Esta construcción es una contribución a la comunidad académica y un plus de esta investigación. Este nuevo conjunto de datos fue entrenado, probado y validado para que clasifique de acuerdo con el Sistema Internacional de líquidos corporales mediante el análisis de imágenes de citología de líquido pleural utilizados para el diagnóstico de células neoplásicas. En la Tabla 8, observamos la estructura del conjunto de datos.

Tabla 8. Estructura del conjunto de datos

Area de datos	Salud	Tipo de datos	Imagen
Formato de datos	PNG	Fuente de datos	HNCH
Tipo de etiquetado	Clasificación		
Utilización de datos	Centros de salud de asistencia al diagnóstico para brindar apoyo a los médicos en el diagnóstico de células cancerosas.	Año de construcción de datos/Cantidad de construcción de datos	2024 2640 imágenes de portaobjetos completos - WSI

5.5. Cumplimiento de normativas éticas y legales

Se realizaron los trámites correspondientes ante el Comité Institucional de Ética en la Investigación del Hospital Nacional Cayetano Heredia, obteniendo la autorización inscrita con el código 102-2024 de fecha 12/12/2024.

CAPÍTULO 6. DESARROLLO DEL MODELO

Se describe la metodología para elaborar el artefacto en este caso el algoritmo basado en aprendizaje profundo para clasificar células neoplásicas en imágenes citológicas de líquido pleural. Esta metodología está enmarcada en un proyecto tipo práctico y experimental, con el propósito de solucionar problemas y contribuir al desarrollo del conocimiento científico, el cual es el enfoque de la DSR detallado en el Capítulo 4.

Se considera práctico, porque se trata de la implementación de un modelo real para solucionar un problema específico: el diagnóstico de células cancerosas en imágenes de líquido pleural, contribuyendo así a un diagnóstico anticipado. También se califica como experimental ya que se realizarán experimentos y pruebas durante el proceso de implementación. Incluye el entrenamiento, ajuste y validación del modelo, junto a la generación y evaluación de datos y ajustes de acuerdo con resultados obtenidos para mejorar el rendimiento del modelo.

Este desarrollo se estructura en etapas donde sobresale la etapa de la construcción del conjunto de datos (Capítulo 5) como una de las partes más importantes de esta investigación ya que la calidad y cantidad de los datos de entrada afectará de gran manera en el desempeño del modelo.

El modelo de aprendizaje profundo propuesto en esta investigación se fundamenta en un flujo metodológico estructurado que inicia con la construcción del conjunto de datos de citología de líquido pleural, tal como se muestra en la Figura 29.

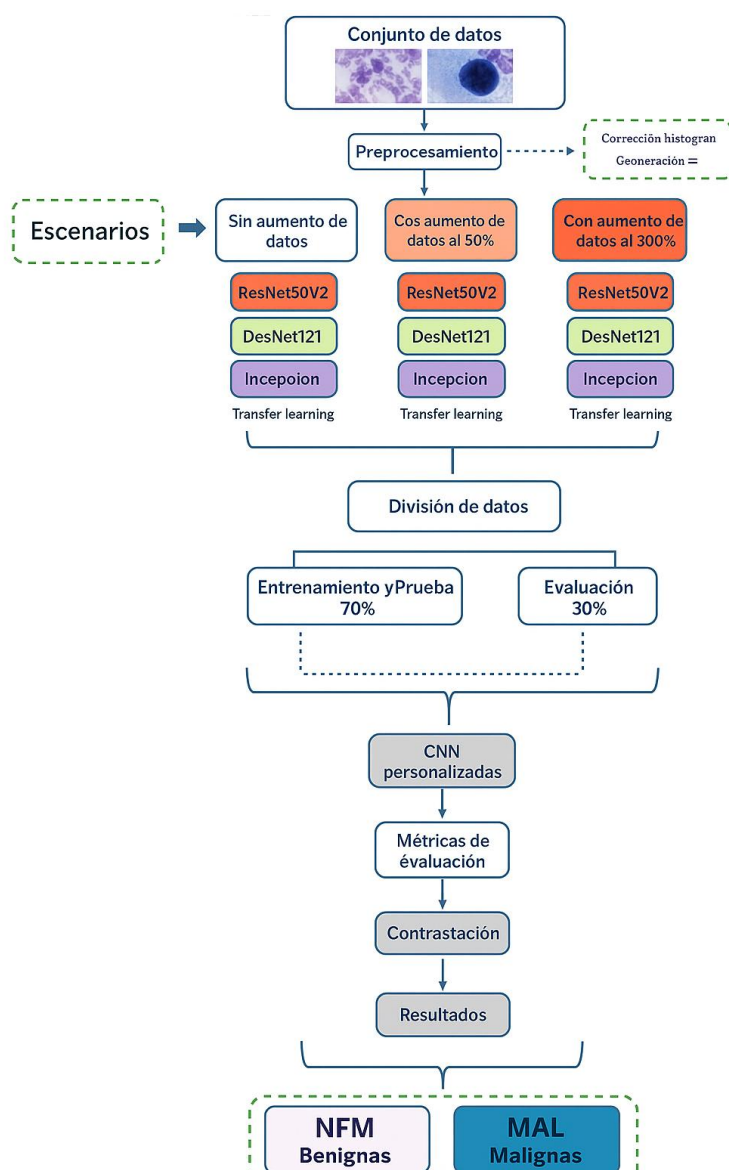


Figura 29. *Modelado propuesto*

6.1. Desarrollo de experimentos

Los experimentos de esta investigación fueron realizados en Jupyter ya que proporciona un entorno interactivo donde se puede escribir y ejecutar los códigos y visualizar datos. Además, se utilizaron las herramientas de desarrollo como el lenguaje Python y algunas principales bibliotecas: Tensor Flow, Keras, OpenCV.

En el capítulo 5 se describió la construcción de un conjunto de datos con imágenes de citología de líquido pleural que contiene 2640 imágenes. Para esta parte de la investigación se tomó en cuenta dos razones; la primera fue la opinión de los expertos respecto a las disparidades del número de imágenes por categoría diagnóstica y la segunda que se obtuvo un conjunto de datos externo para hacer la validación y era de clasificación binaria, por lo que se optó por hacer una salida binaria entre las categorías diagnóstica MAL y NFM seleccionando 1292 imágenes para los experimentos del modelado. En la Tabla 9 se observa las categorías diagnósticas a trabajar en esta investigación.

Tabla 9. Categorías diagnósticas para la investigación.

SISTEMA INTERNACIONAL PARA INFORMAR CITOLOGIA DE LIQUIDOS SEROSOS: RIESGO DE MALIGNIDAD	Categoría	Total de imágenes
MALIGNO	MAL	645
NEGATIVA PARA MALIGNIDAD	NFM	647
TOTAL		1292

6.2. Generación de nuevas imágenes

En la investigación se implementó un modelo de data augmentation geométrico y fotométrico, utilizando la clase ImageDataGenerator de TensorFlow de Python. Esta técnica de aumentar datos tiene como objetivo incrementar la variedad del conjunto de entrenamiento y mejorar la capacidad de generalización del modelo.

Se aplicaron e implementaron transformaciones que siguen un enfoque de augmentation estocástico en tiempo real, donde cada imagen del conjunto de entrenamiento fue sometida aleatoriamente a las siguientes modificaciones:

1. Normalización de Píxeles. - Rescaling (1/255): Normalización de los valores de píxeles del rango [0, 255] al rango [0, 1], facilitando la convergencia del modelo y estabilizando el proceso de entrenamiento.
2. Transformaciones Geométricas. - Horizontal Flip (True): Reflexión horizontal aleatoria de las imágenes, duplicando efectivamente la variabilidad de poses y orientaciones en el conjunto de datos.

Vertical Flip (True): Reflexión vertical aleatoria, proporcionando robustez adicional ante diferentes orientaciones de los objetos de interés.

Rotation Range (30°): Rotación aleatoria de las imágenes en un rango de ± 30 grados, simulando diferentes ángulos de captura y mejorando la invarianza rotacional del modelo.

3. Transformaciones de Desplazamiento. - Width Shift Range (0.15): Desplazamiento horizontal aleatorio del 15% del ancho total de la imagen, simulando variaciones en la posición del objeto dentro del encuadre.

Height Shift Range (0.15): Desplazamiento vertical aleatorio del 15% de la altura total de la imagen, proporcionando robustez ante variaciones en la posición vertical del objeto.

4. Transformaciones de Escala. - Zoom Range (0.15): Aplicación aleatoria de zoom con un factor de variación del 15%, simulando diferentes distancias de captura y tamaños de objeto en la imagen.

En la Figura 30 se visualiza los diferentes parámetros utilizados en la transformación de imágenes usando el aumento de datos con la técnica data augmentation.

```

train_datagen = ImageDataGenerator(
    rescale=1/255.,
    horizontal_flip=True,
    vertical_flip=True,
    rotation_range=30,
    width_shift_range=0.15,
    height_shift_range=0.15,
    zoom_range=0.15
)

```

Figura 30. Parámetros del ImageDataGenerator para la transformación de imágenes.

Se aplica a cada clase: giro horizontal, giro vertical, giro por ángulo, rango de cambio de altura, rango de cambio de ancho y zoom 0.15. Posteriormente se guardaron las imágenes generadas y se unieron con las imágenes originales normales para conformar un solo conjunto de datos. En la Figura 31 se muestra los efectos aplicados.

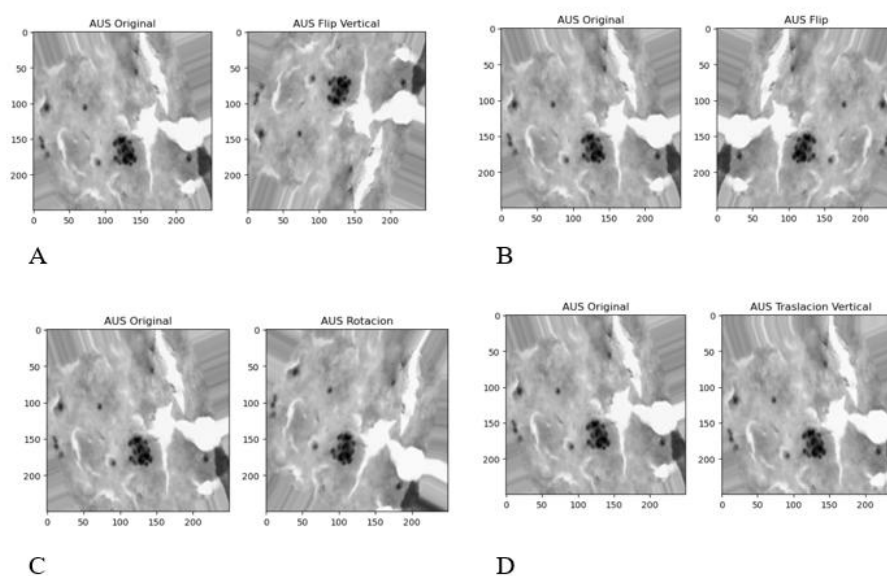


Figura 31. Efectos aplicados en: A vertical_flip=True, B horizontal_flip=True, C rotation_range=30 y D width_shift_range=0.15.

Justificación Técnica

La selección de estos parámetros se fundamenta en los siguientes criterios:

- **Preservación de la Integridad Semántica:** Los rangos de transformación (15% para desplazamientos y zoom, 30° para rotación) fueron seleccionados para mantener la interpretabilidad visual y la coherencia semántica de las imágenes.
- **Aumento de Variabilidad sin Distorsión:** Los valores moderados evitan la generación de imágenes artificiales que puedan introducir ruido en el proceso de aprendizaje.
- **Mejora de la Robustez:** Las transformaciones geométricas combinadas proporcionan invarianza ante traslaciones, rotaciones y cambios de escala, características deseables en tareas de clasificación/detección de imágenes.

Impacto en el Rendimiento del Modelo

La técnica de data augmentation posibilita:

- Expansión del conjunto de entrenamiento sin necesidad de recolección adicional de datos
- Reducción del sobreajuste mediante la introducción de variabilidad controlada
- Mejora de la generalización del modelo ante nuevas instancias no vistas durante el entrenamiento
- Optimización de recursos computacionales al generar variaciones en tiempo real durante el entrenamiento
- La implementación de estas técnicas de data augmentation constituye una práctica estándar en el campo del aprendizaje profundo para visión por computadora, respaldada por evidencia empírica en la literatura científica especializada.

6.3. Creación de escenarios

Luego se trabajó en el modelado de las tres redes neuronales convolucionales CNN utilizando como cabecera de la red a arquitecturas de aprendizaje por transferencia: ResNet50V2, DesNet121 e Inceptionv3, para posteriormente entrenar, validar y evaluar (prueba). Para evitar el riesgo de sobreajuste se realizó la debida personalización de las capas finales de los modelos para clasificar el conjunto de datos. Se usaron pesos entrenados previamente del conjunto de datos ImageNet conocida como Transfer Learning.

Se crearon tres escenarios para obtener como salida una clasificación binaria de NFM (benignas) y MAL (malignas):

- **Escenario 1:** Conjunto de datos con 1292 imágenes citológicas sin aumento de datos.
- **Escenario 2:** Conjunto de datos con 1292 imágenes citológicas con aumento de datos al 50%, total 1938 imágenes.
- **Escenario 3:** Conjunto de datos con 1292 imágenes citológicas con aumento de datos al 300%, total 3876 imágenes.

Para la etapa de evaluación, se utilizó matriz de rendimiento como la matriz de precisión.

En el Capítulo 1 se presentó el modelado de la propuesta de esta investigación.

6.4. Experimentos

Los experimentos se realizaron con dos conjuntos de datos:

Primer conjunto de datos: Creado por la investigadora en colaboración con el Departamento de Anatomía Patológica del Hospital Nacional Cayetano Heredia (HNCH), cuenta con la autorización para la realización de la investigación del Comité de Ética del HNCH. El experimento se realizó con 1292 imágenes citológicas de líquido pleural.

Especialistas patólogos con más de 10 años de experiencia anotaron cada imagen en la categoría diagnóstica correspondiente: 645 imágenes MAL (malignas) y 647 imágenes NFM (benignas). Este conjunto de datos fue utilizado para el entrenamiento, validación y evaluación del modelo. La categoría diagnóstica y el número de imágenes se visualiza en la Tabla 10.

Tabla 10. Categoría diagnóstica y número de imágenes del conjunto de datos HNCH

Categoría	MAL	NFM	Total
Nro de Imágenes	645	647	1292
Descripción	Malignas	Benignas	

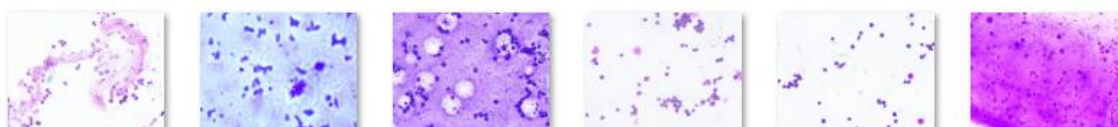
Segundo conjunto de datos: Obtenido a través del repositorio Kaggle, conjunto de datos de nombre *Body cavity fluid cytology images* (Kaggle & Parikshit Sanyal, 2021), que consta 533 imágenes MAL (malignas) y 160 imágenes NFM (benignas), total de 693 imágenes. Estas imágenes fueron tomadas con un aumento de 40X y luego se redimensionaron a 250 x 250 píxeles. Este conjunto de datos fue utilizado para la contrastación del modelo. La categoría diagnóstica y el número de imágenes se visualiza en la Tabla 11.

Tabla 11. Categoría diagnóstica y número de imágenes del conjunto de datos KAGGLE

Categoría	MAL	NFM	Total
Nro de Imágenes	533	160	693
Descripción	Malignas	Benignas	

En la Figura 32 se muestra imágenes del conjunto de datos *Body cavity fluid cytology images*, tomado de (Kaggle & Parikshit Sanyal, 2021), conformado por muestras NFM-Negativa para malignidad y MAL-Malignas.

a) NFM - Negativa para malignidad



b) MAL - Malignas

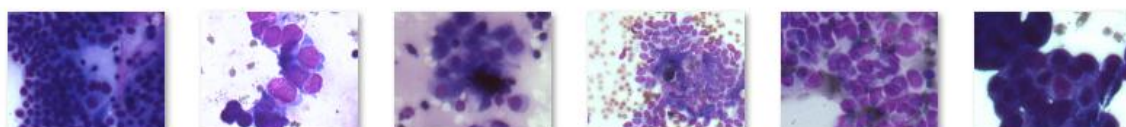


Figura 32. *Imágenes del conjunto de datos Body cavity fluid cytology images: a) NFM-Negativa para malignidad y b) MAL-Malignas.*

6.5. Modelado para clasificar imágenes

Extracción de características: También conocido como representación y descripción. En esta etapa, a la imagen se le extraen un conjunto de características que describen las propiedades físicas de los objetos como son: el color, textura, área, perímetro, ancho, alto, promedio de intensidad, centroide, orientación, color, entre otros (Silva, 2019).

Luego de la extracción de características significativas del conjunto de datos, que contenía células NFM (benignas) y MAL (malignas), se dividió en datos de entrenamiento 70% y validación 30%.

Se crearon tres redes neuronales convolucionales CNN que tenían como primera capa las arquitecturas ResNet50V2, DesNet121 e InceptionV3 respectivamente. Las capas subsiguientes están ordenadas como se describe a continuación: Capa de agrupación de la media total, capa densa, capa dropout, capa de ruido gaussiano, capa densa, capa dropout y capa densa (salida).

Usamos ResNet50V2, que es una arquitectura avanzada pre entrenada con datos de ImageNet, como base para extraer características importantes de las imágenes. Al inicio se congela el modelo base para que no se modifique durante el entrenamiento. Luego, se añaden nuevas capas para personalizar la red: una capa que reduce la información, dos capas densas que procesan los datos y una capa final que da las probabilidades de pertenecer a cada clase. Para evitar errores y mejorar el rendimiento, se usan técnicas como Dropout (apaga neuronas al azar) y GaussianNoise (añade ruido) durante el entrenamiento. Al final, el modelo es capaz de adaptarse a un problema específico de clasificación. Se realiza el mismo procedimiento para las arquitecturas DesNet121 e Inceptionv3. En la Figura 33 se visualiza la arquitectura de cada red neuronal.

Para mejorar el desempeño del modelo, se implementó un clasificador de votación que combina las predicciones de las tres arquitecturas (ResNet, DenseNet e Inception). Se exploraron dos estrategias de combinación: hard voting, se selecciona la clase mayoritaria entre las predicciones individuales, y soft voting, se promedian las probabilidades de las predicciones y se selecciona la clase con mayor probabilidad promedio.

El clasificador de votación permitió unir las diferentes combinaciones de los tres clasificadores ResNet50V2, DesNet121 e InceptionV3 (ResNet-DesNet, ResNet-Inception, DesNet-Inception y ResNet-DesNet-Inception).

a)

Layer (Type)	Output Shape	Param #
resnet50v2 (Functional)	(None, 8, 8, 2048)	23,564,800
global_average_pooling2d_1 (GlobalAveragePooling2D)	(None, 2048)	0
dense_1 (Dense)	(None, 512)	1,049,088
dropout_1 (Dropout)	(None, 512)	0
gaussian_noise (GaussianNoise)	(None, 512)	0
dense_2 (Dense)	(None, 128)	65,664
dropout_2 (Dropout)	(None, 128)	0
dense_3 (Dense)	(None, 2)	258
Total params	24,679,810	
Trainable params	1,115,010	
Non-trainable params	23,564,800	

b)

Layer (Type)	Output Shape	Param #
densenet121 (Functional)	(None, 7, 7, 1024)	7,037,504
global_average_pooling2d_2 (GlobalAveragePooling2D)	(None, 1024)	0
dense_4 (Dense)	(None, 512)	524,800
dropout_3 (Dropout)	(None, 512)	0
gaussian_noise_1 (Gaussian- Noise)	(None, 512)	0
dense_5 (Dense)	(None, 128)	65,664
dropout_4 (Dropout)	(None, 128)	0
dense_6 (Dense)	(None, 2)	258
Total params	7,628,226	
Trainable params	590,722	
Non-trainable params	7,037,504	

c)

Layer (Type)	Output Shape	Param #
inception_v3 (Functional)	(None, 6, 6, 2048)	21,802,784
global_average_pooling2d_3 (Glo- balAveragePooling2D)	(None, 2048)	0
dense_7 (Dense)	(None, 512)	1,049,088
dropout_5 (Dropout)	(None, 512)	0
gaussian_noise_2 (GaussianNoise)	(None, 512)	0
dense_8 (Dense)	(None, 128)	65,664
dropout_6 (Dropout)	(None, 128)	0
dense_9 (Dense)	(None, 2)	258
Total params	22,917,794	
Trainable params	1,115,010	
Non-trainable params	21,802,784	

Figura 33. Arquitectura de red neuronal convolucional utilizando a) RedNet50V2, b)DesNet131 y c) InceptionV3

6.6. Distribución del conjunto de datos

Para la etapa de entrenamiento y validación se realizó una división del conjunto de datos en las proporciones de 70% para entrenamiento y validación y 30% para evaluación. Se crearon tres escenarios para comparar los resultados y elegir el que obtuviera mayor rendimiento:

Escenario 1: Sin generación de imágenes nuevas

En este escenario fueron 1292 imágenes, como se muestra en la Tabla 12.

Tabla 12. Número de imágenes al inicio del experimento

Categoría	MAL	NFM	Total
	645	647	1292
70%	452	453	905
30%	193	194	387

Escenario 2: Generación de imágenes nuevas al 50%

En este escenario fueron 1292 imágenes, luego de aplicar aumento de datos se obtuvo 1938 imágenes, como se muestra en la Tabla 13.

Tabla 13. Número de imágenes generadas al 50%

Categoría	MAL	NFM	Total
	645	647	1292
50%	323	324	646

Con aumento	968	971	1938
70%	678	680	1378
30%	290	291	581

Escenario 3: Generación de imágenes nuevas al 300%

En este escenario fueron 1292 imágenes, luego de aplicar aumento de datos se obtuvo 5168 imágenes, como se muestra en la Tabla 14.

Tabla 14. Número de imágenes generadas al 300%

Categoría	MAL	NFM	Total
	645	647	1292
300%	1935	1941	3876
Con aumento	2580	2588	5168
70%	1806	1812	3618
30%	774	776	1550

6.7. Entrenamiento y validación

Escenario 1: Sin generación de imágenes nuevas

En el escenario 1 se trabaja el conjunto de datos sin generación de nuevas imágenes, configurando los parámetros para cada arquitectura y obteniendo los resultados del proceso de entrenamiento y validación según se muestra en la Tabla 15.

Tabla 15. Parámetros y resultados obtenidos en el entrenamiento y validación del escenario 1.

Categoría	Característica	ResNet	DesNet	Inception
Parámetros	Epocas	82/100	100/100	77/100
	# pasos	30	20	20
	Tiempo x paso	4/70seg	3/57seg	2/32seg
	Learning Rate	2.00E-05	1.00E-04	2.00E-05
Resultados	Loss	0.1167	0.2316	0.1054
	Accuracy	0.9561	0.9058	0.9626
	Val loss	0.1434	0.1631	0.1280
	Val accuracy	0.9302	0.9302	0.9535

ResNet50V2 logra un accuracy de 0.9561 y un val_accuracy de 0.9302 indicando una buena capacidad de generalización a los datos de validación y los valores de pérdida indican un error bajo tanto en entrenamiento como en validación.

En DesNet121 los valores de accuracy 0.9058 y val_Accuracy 0.9302 tienen una mejora de 0.0244 lo que refleja una sólida capacidad de generalización a los datos de validación.

En el modelo Inception se obtiene un accuracy sobresaliente de 0.9626 y un val_accuracy de 0.9535 lo que indica un excelente ajuste al conjunto de datos, los valores de loss y val_loss son los más bajos de las tres arquitecturas lo que confirma un error muy reducido en ambos procesos.

En la figura 34 se muestran los gráficos provenientes de la precisión y pérdida en el entrenamiento y la validación del escenario 1. La línea de color verde corresponde a la precisión y la de color rojo al valor de pérdida del modelo ResNet50V2 en el escenario 1.

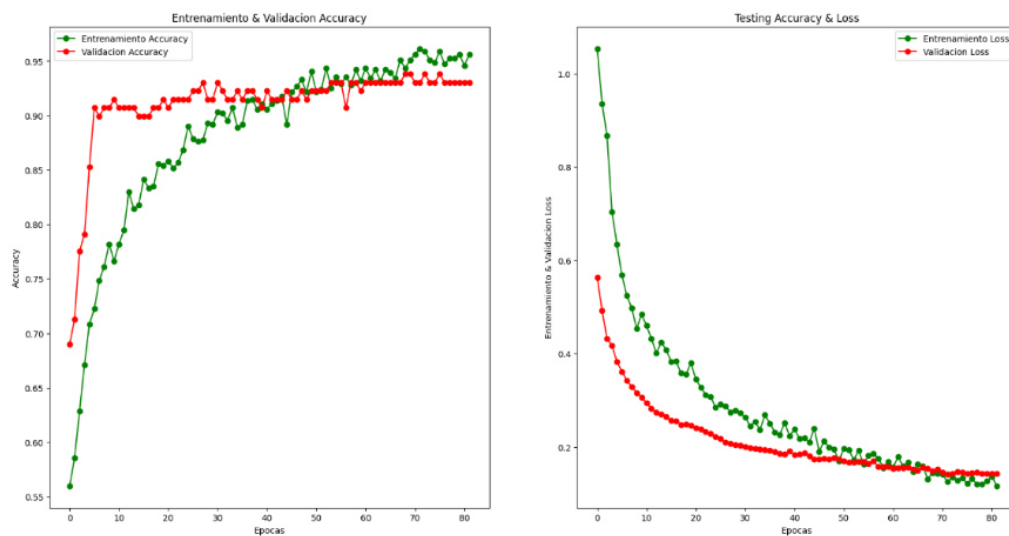


Figura 34. Resultado del entrenamiento y validación para ResNet50V2 escenario1

En la Figura 35 se muestran los gráficos provenientes de la precisión y pérdida en el entrenamiento y la validación del escenario 1. La línea de color verde corresponde a la precisión y la de color rojo al valor de pérdida del modelo DesNet121 en el escenario 1.

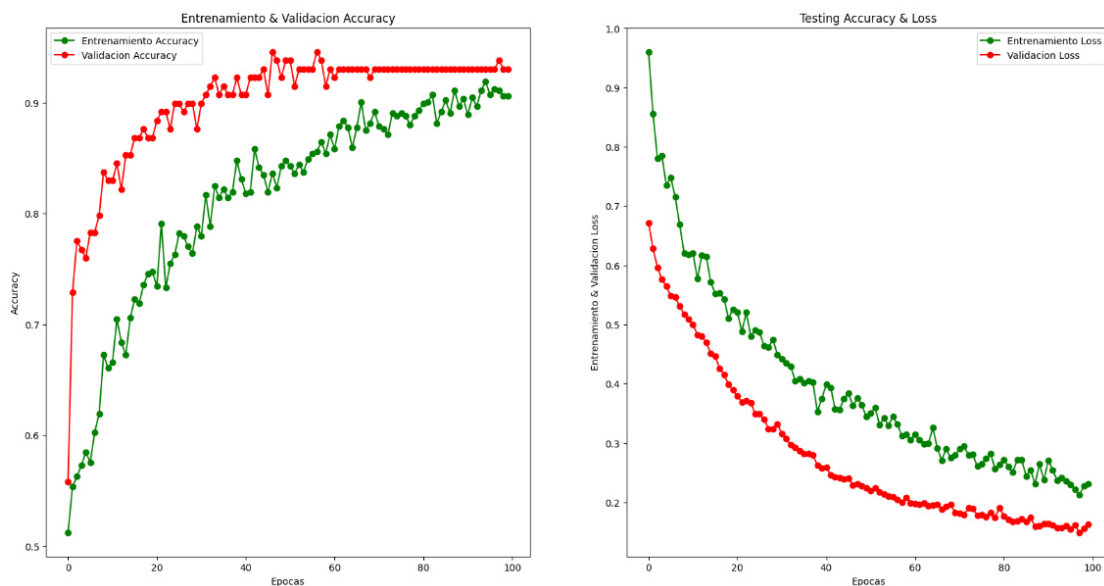


Figura 35. Resultado del entrenamiento y validación para Desnet121 escenario1

En la Figura 36 se muestran los gráficos provenientes de la precisión y pérdida en el entrenamiento y la validación del escenario 1. La línea de color verde corresponde a la precisión y la de color rojo al valor de pérdida del modelo Inception en el escenario 1.

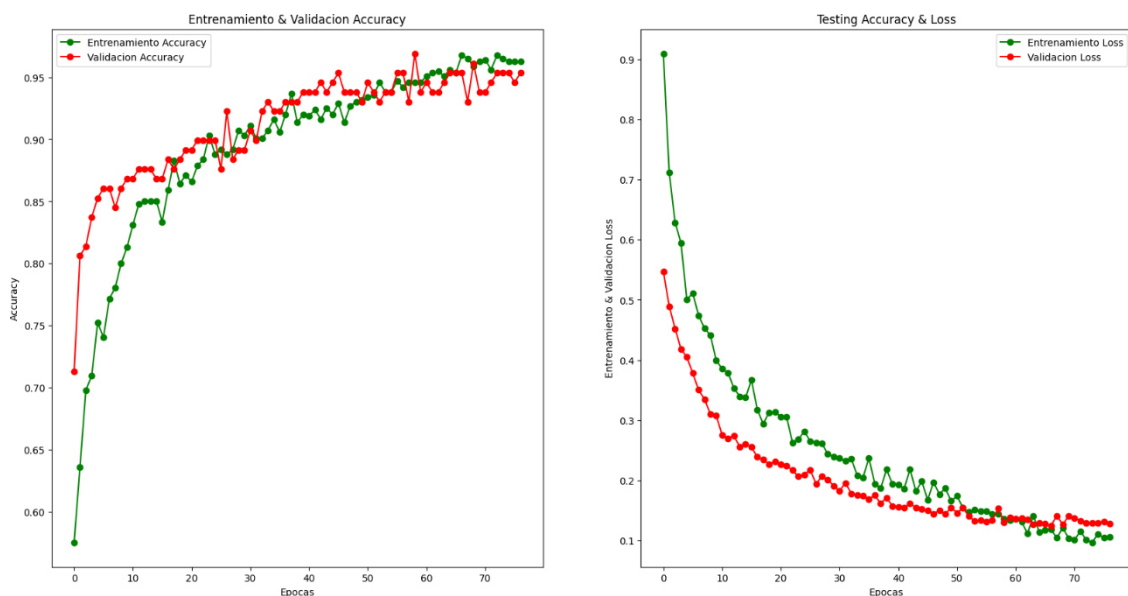


Figura 36. Resultado del entrenamiento y validación para Inception escenario1

Escenario 2: Generación de imágenes nuevas al 50%

En el escenario 2 se trabaja el conjunto de datos con generación de nuevas imágenes al 50%, configurando los parámetros para cada arquitectura y obteniendo los resultados del proceso de entrenamiento y validación según se muestra en la Tabla 16.

Tabla 16. Parámetros y resultados obtenidos en el entrenamiento y validación del escenario 2

Categoría	Característica	ResNet	DesNet	Inception
Parámetros	Epocas	55/100	89/100	51/100
	# pasos	30	30	30
	Tiempo x paso	3/83seg	3/90seg	2/51seg
	Learning Rate	2.00E-05	4.00E-06	2.00E-05
Resultados	Loss	0.1785	0.3039	0.1765
	Accuracy	0.9296	0.8625	0.939
	Val loss	0.2605	0.2512	0.2852
	Val accuracy	0.8	0.9175	0.8711

El modelo ResNet50V2 muestra buen desempeño con alta precisión en los datos de entrenamiento (92.96%) y validación (80%). Sin embargo, el aumento en la pérdida de validación (mayor que la pérdida de entrenamiento) podría ser indicativo de un leve sobreajuste, lo cual sería normal en esta etapa del entrenamiento.

El modelo DesNet121 muestra relativa precisión en los datos de entrenamiento (86.25%) y validación (91.75%). Se observa que la pérdida de validación es menor que la pérdida de entrenamiento, lo que indica que el conjunto de datos de validación puede ser más fácil de predecir para el modelo que el conjunto de datos de entrenamiento.

Inception logra el mejor accuracy de 93.90%, de las tres arquitecturas, demostrando un excelente ajuste a los datos de entrenamiento, sin embargo, el Val_accuracy decae a 87.11% lo que refleja una ligera tendencia al sobreajuste, con respecto a los valores de loss y val-loss el error de validación es mucho más elevado que las otras arquitecturas, lo que podría indicar retos para la generalización en el modelo Inception.

En la Figura 37 se muestran los gráficos provenientes de la precisión y pérdida en el entrenamiento y la validación del escenario 2. La línea de color verde corresponde a la precisión y la de color rojo al valor de pérdida del modelo ResNet50V2 en el escenario 2.

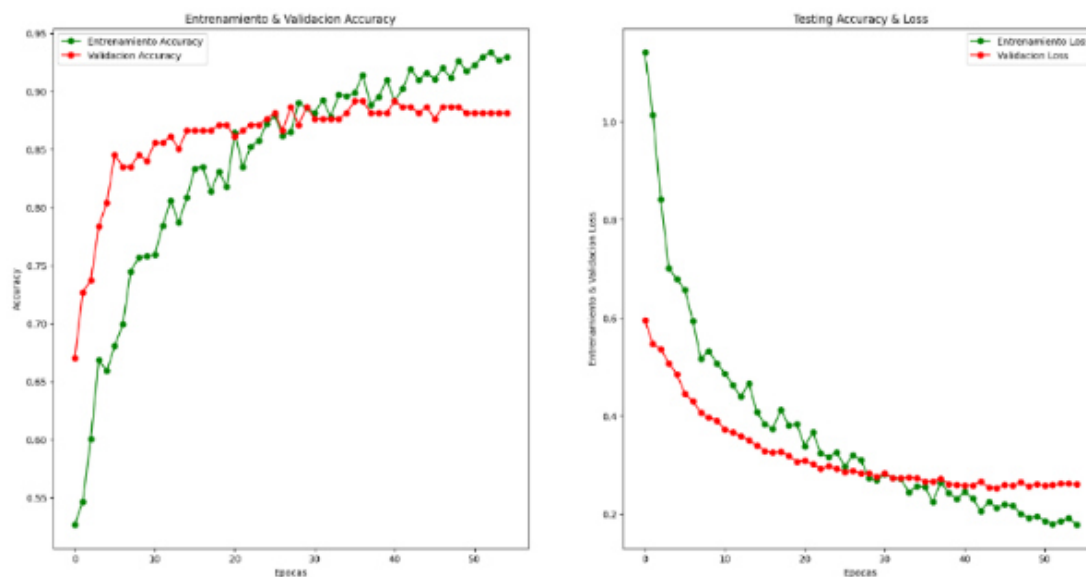


Figura 37. Resultado del entrenamiento y validación para ResNet50V2 escenario 2.

En la Figura 38 se muestran los gráficos provenientes de la precisión y pérdida en el entrenamiento y la validación del escenario 2. La línea de color verde corresponde a la precisión y la de color rojo al valor de pérdida del modelo DesNet121 en el escenario 2.

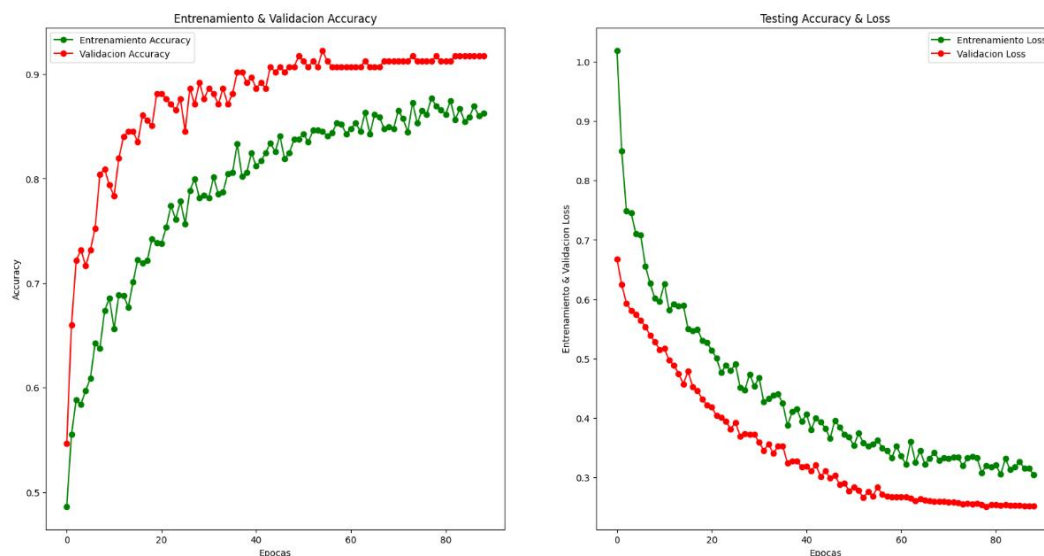


Figura 38. Resultado del entrenamiento y validación para DesNet121 escenario 2

En la Figura 39 se muestran los gráficos provenientes de la precisión y pérdida en el entrenamiento y la validación del escenario 2. La línea de color verde corresponde a la precisión y la de color rojo al valor de pérdida del modelo Inception en el escenario 2.

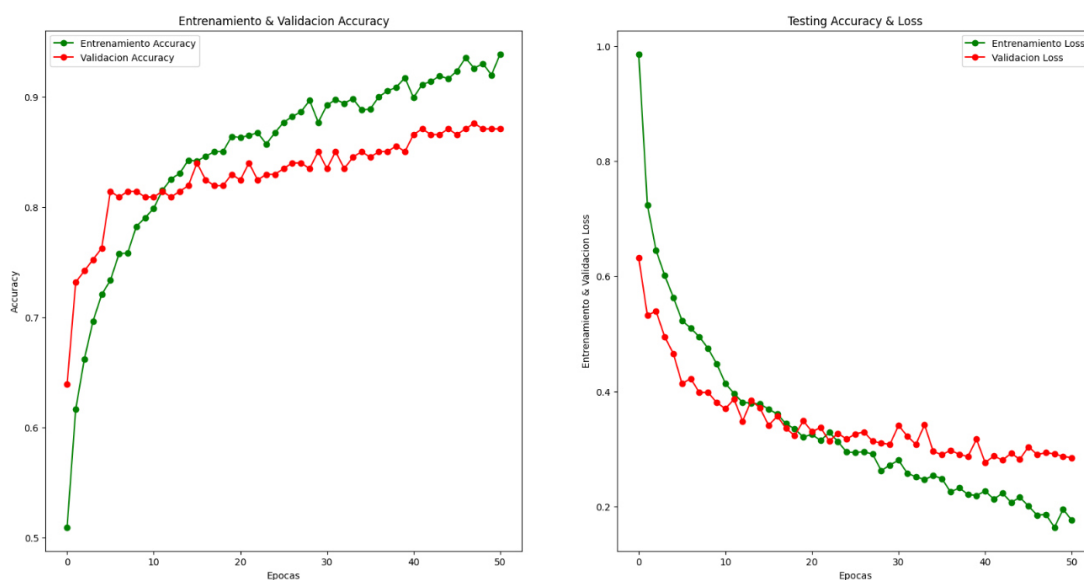


Figura 39. Resultado del entrenamiento y validación para Inception escenario 2

Escenario 3: Generación de imágenes nuevas al 300%

En el escenario 3 se trabaja el conjunto de datos con generación de nuevas imágenes al 300%, configurando los parámetros para cada arquitectura y obteniendo los resultados del proceso de entrenamiento y validación según se muestra en la Tabla 17.

Tabla 17. Parámetros y resultados obtenidos en el entrenamiento y validación del escenario 3

Categoría	Característica	ResNet	DesNet	Inception
Parámetros	Epocas	52/100	100/100	60/100
	# pasos	78	78	78
	Tiempo x paso	2/182seg	3/227seg	2/129seg
	Learning Rate	4.00E-06	2.00E-05	4.00E-06
Resultados	Loss	0.1015	0.1412	0.1013
	Accuracy	0.9584	0.9423	0.9635
	Val loss	0.1681	0.1884	0.1668
	Val accuracy	0.9323	0.9188	0.9284

ResNet50V2 alcanza un accuracy de 0.9584 y un val_accuracy de 0.9323 demostrando una modesta capacidad de generalización a los datos de validación sus valores de pérdida incrementan en la etapa de validación, sin embargo, refleja un bajo error en ambas etapas, lo que sugiere que el modelo presenta un buen ajuste.

DesNet121 obtiene un accuracy de 0.9423 y un val_accuracy de 0.9188 disminuyendo su precisión en la etapa de validación, sus valores de pérdida son ligeramente más altos, lo que indica que existe un margen de mejora para ajustar el modelo.

Inception logra un accuracy de 0.9635 y un val_accuracy de 0.9284 lo que sugiere un modesto desempeño en la validación, los valores de pérdida son bajos en ambas etapas, destacando la eficiencia del modelo.

En la Figura 40 se muestran los gráficos provenientes de la precisión y pérdida en el entrenamiento y la validación del escenario 3. La línea de color verde corresponde a la precisión y la de color rojo al valor de pérdida del modelo ResNet50V2 en el escenario 3.

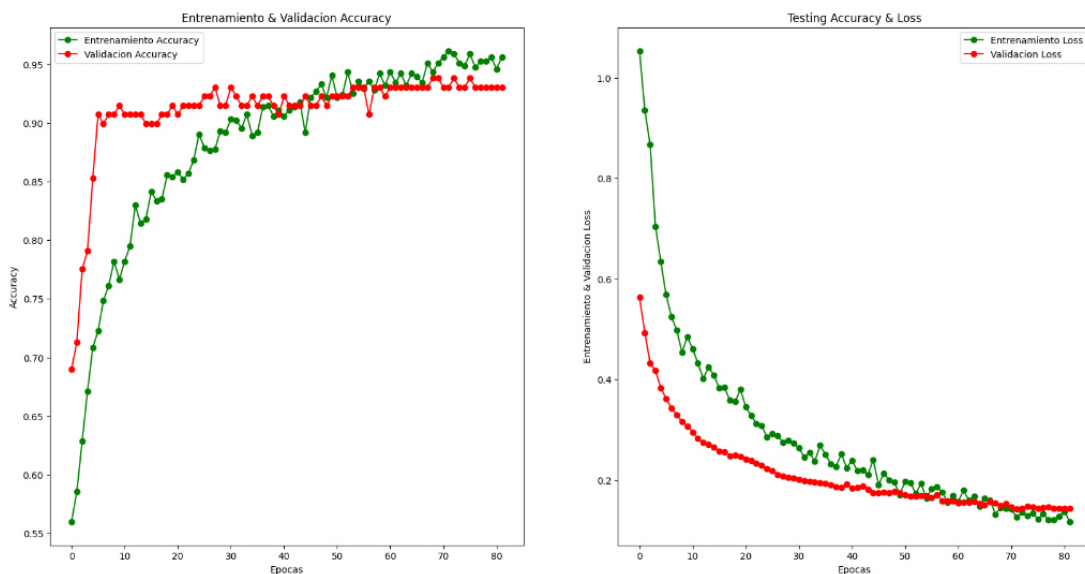


Figura 40. Resultado del entrenamiento y validación para ResNet50V2 escenario 3.

En la Figura 41 se muestran los gráficos provenientes de la precisión y pérdida en el entrenamiento y la validación del escenario 3. La línea de color verde corresponde a la precisión y la de color rojo al valor de pérdida del modelo DesNet121 en el escenario 3.

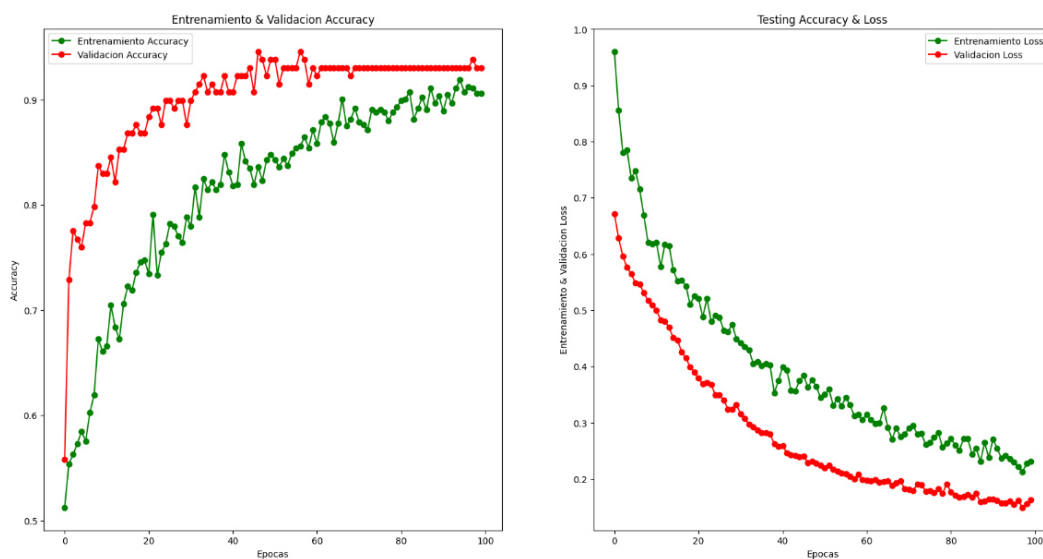


Figura 41. Resultado del entrenamiento y validación para DesNet121 escenario 3

En la Figura 42 se muestran los gráficos provenientes de la precisión y pérdida en el

entrenamiento y la validación del escenario 3. La línea de color verde corresponde a la precisión y la de color rojo al valor de pérdida del modelo Inception en el escenario 3.

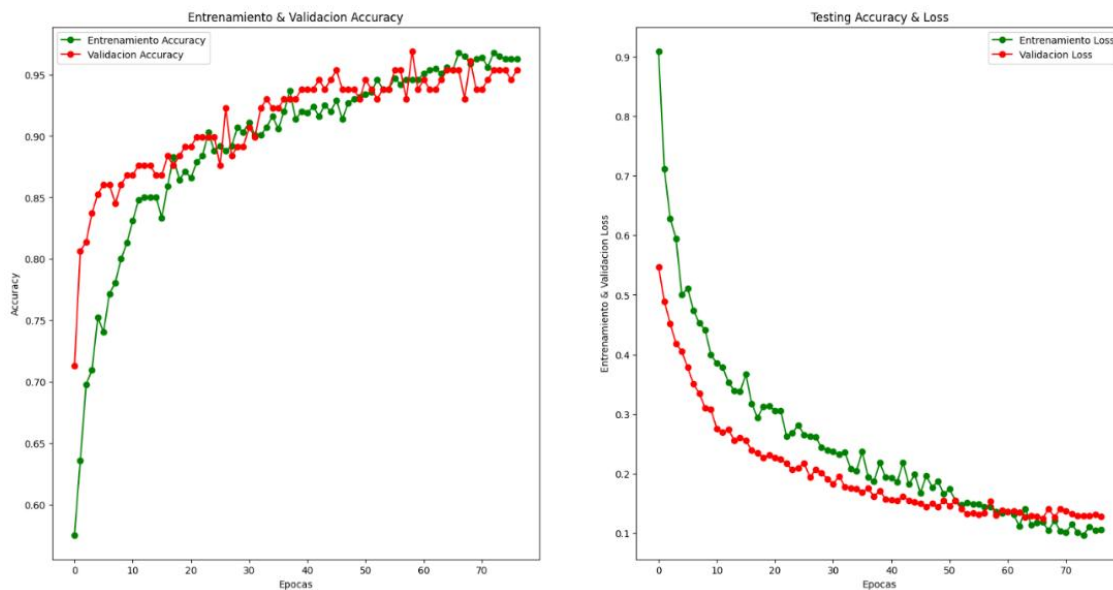


Figura 42. Resultado del entrenamiento y validación para Inception escenario 3

CAPÍTULO 7. RESULTADOS Y DISCUSIÓN

De acuerdo con la metodología DSR, desarrollamos la actividad de evaluación que se refiere al desarrollo de indicadores por los cuales se mide el rendimiento del artefacto.

7.1. Evaluación por escenarios

Escenario 1: Sin generación de imágenes nuevas

En el escenario 1 se trabajó el conjunto de datos con las imágenes originales sin generación de imágenes nuevas, obteniendo los resultados en la etapa de evaluación (testing) de las tres arquitecturas, conforme se visualiza en la Figura 43.

A)					B)					C)				
13/13 [*****]-21s 2s/step					13/13 [*****]-27s 2s/step					13/13 [*****]-16 1s/step				
Classification Report:					Classification Report:					Classification Report:				
	precision	recall	f1-score	support		precision	recall	f1-score	support		precision	recall	f1-score	support
MAL	0.94	0.91	0.92	207	MAL	0.94	0.86	0.89	207	MAL	0.95	0.93	0.94	207
NFM	0.90	0.93	0.92	181	NFM	0.85	0.93	0.89	181	NFM	0.92	0.94	0.93	181
accuracy			0.92	388	accuracy			0.89	388	accuracy			0.94	388
macro avg	0.92	0.92	0.92	388	macro avg	0.89	0.89	0.89	388	macro avg	0.93	0.94	0.94	388
weighted avg	0.92	0.92	0.92	388	weighted avg	0.90	0.89	0.89	388	weighted avg	0.94	0.94	0.94	388

Figura 43. Evaluación escenario 1: (A) ResNet50V2, (B) DesNet21 y (C) InceptionV3

La arquitectura ResNet50V2 presenta una precisión general del 92%, lo que indica un buen desempeño en la clasificación, los valores de precisión recall f1-score son consistentes en las dos categorías de salida con una media ponderada de 0.92 para cada métrica, lo que refleja un modelo equilibrado en cuanto a la minimización de falsos positivos y falsos negativos.

DesNet121 muestra un accuracy del 89% ligeramente inferior en comparación con sus métricas de precisión y recall, sugiriendo que el modelo tiene un menor desempeño en

cuanto a la generalización, pero, aun es adecuado para las tareas de clasificación con un f1-score consistente en ambas clases.

La arquitectura Inception alcanza un accuracy del 93% y una media ponderada de 0.94 en las medidas de precisión, recall y f1-score, lo que evidencia un modelo bien ajustado y altamente eficaz en la clasificación con una excelente capacidad de identificación de imágenes en ambas categorías diagnósticas.

Escenario 2: Generación de imágenes nuevas al 50%

En el escenario 2 se trabajó el conjunto de datos con las imágenes generadas al 50%, obteniendo los resultados en la etapa de evaluación (testing) de las tres arquitecturas, conforme se visualiza en la Figura 44.

A)					B)					C)				
19/19 [*****]-38s 2s/step					19/19 [*****]-41s 2s/step					19/19 [*****]-236 1s/step				
Classification Report:					Classification Report:					Classification Report:				
	precision	recall	f1-score	support		precision	recall	f1-score	support		precision	recall	f1-score	support
MAL	0.93	0.92	0.92	284	MAL	0.91	0.83	0.87	284	MAL	0.92	0.87	0.89	284
NFM	0.92	0.94	0.93	298	NFM	0.85	0.93	0.89	298	NFM	0.88	0.93	0.90	298
accuracy			0.93	582	accuracy			0.88	582	accuracy			0.90	582
macro avg	0.93	0.93	0.93	582	macro avg	0.88	0.88	0.88	582	macro avg	0.90	0.90	0.90	582
weighted avg	0.93	0.93	0.93	582	weighted avg	0.88	0.88	0.88	582	weighted avg	0.90	0.90	0.90	582

Figura 44. Evaluación escenario 2: (A) ResNet50V2, (B) DesNet21 y (C) InceptionV3

La arquitectura Resnet50V2 logra un accuracy del 93% y métricas de precision, recall y f1-score con valores promedios ponderados de 0.93, lo que indica una excelente capacidad para identificar ambas categorías diagnosticas con una mínima repercusión de error.

Por su lado Desnet121 alcanza un accuracy del 88% junto con valores promedios ponderados de precisión, recall y f1-score de 0.88, sugiriendo que el modelo tiene un desempeño adecuado, aunque menos efectivo que los otros modelos.

La arquitectura Inception presenta un accuracy del 90% con valores promedios ponderados de precisión, recall y f1-score de de 0.90 reflejando un modelo con desempeño equilibrado y una modesta capacidad para distinguir ambas categorías diagnosticas.

Escenario 3: Generación de imágenes nuevas al 300%

En el escenario 3 se trabajó el conjunto de datos con las imágenes generadas al 300%, obteniendo los resultados en la etapa de evaluación (testing) de las tres arquitecturas, conforme se visualiza en la Figura 45.

A)					B)					C)				
49/49 [*****]-79s 2s/step					49/49 [*****]-100s 2s/step					49/49 [*****]-57s 1s/step				
Classification Report:					Classification Report:					Classification Report:				
	precision	recall	f1-score	support		precision	recall	f1-score	support		precision	recall	f1-score	support
MAL	0.95	0.95	0.95	774	MAL	0.95	0.93	0.94	774	MAL	0.96	0.92	0.94	774
NFM	0.95	0.95	0.95	777	NFM	0.94	0.95	0.94	777	NFM	0.95	0.96	0.94	777
accuracy			0.95	1551	accuracy			0.94	1551	accuracy			0.94	1551
macro avg	0.95	0.95	0.95	1551	macro avg	0.94	0.94	0.94	1551	macro avg	0.94	0.94	0.94	1551
weighted avg	0.95	0.95	0.95	1551	weighted avg	0.94	0.94	0.94	1551	weighted avg	0.94	0.94	0.94	1551

Figura 45. Evaluación del escenario 3: (A) ResNet50V2, (B) DesNet12 y (C) InceptionV3

ResNet50V2 muestra un desempeño destacado con un accuracy del 95%, con métricas de precisión, recall y f1-score con valores promedios ponderados de 0.95 lo que refleja un modelo eficiente y consistente para ambas categorías diagnósticas.

DesNet121 logra un accuracy del 94% con valores promedios ponderados de precisión, recall y f1-score de 0.94, que demuestra un desempeño sólido, manteniendo un balance adecuado en ambas categorías diagnósticas.

Finalmente, la arquitectura Inception presenta igual accuracy del 94% y valores promedios ponderados de precisión, recall y f1-score de 0.94, evidenciando un modelo confiable y bien ajustado.

7.2. Resultados

Se evaluaron los resultados de los tres escenarios y se eligió aquel escenario que cumplía con el objetivo de la investigación. Los escenarios fueron evaluados con el 30% del conjunto de datos del HNCH y como segunda evaluación con el 100% de los datos del conjunto de KAGGLE.

Para esta fase se calculó la métrica de evaluación (precisión) de los modelos ResNet50V2, DesNet121 e Inception tanto en el conjunto de datos de imágenes propias (HNCH) y en el conjunto de datos de imágenes de Kaggle. Posteriormente se aplican las combinaciones posibles de estos tres modelos mediante el clasificador de votación tanto de tipo hard voting como de tipo soft voting: ResNet+Inception, DenseNet+Inception, ResNet+DenseNet, ResNet+DenseNet+Inception. Finalmente, se realiza comparaciones combinadas de los modelos en a los tres escenarios.

Escenario 1 Sin generación de imágenes

En la Tabla 18 se visualizan en la parte izquierda la precisión de los modelos ResNet50V2, DesNet12, InceptionV3 del conjunto de datos propios y del conjunto de datos de Kaggle. Por otro lado, en la parte derecha se visualiza todas las combinaciones posibles de los tres modelos anteriormente mencionados mediante el agrupamiento con el clasificador de votación tanto del tipo hard y del tipo soft voting.

Tabla 18. Métrica de evaluación de los modelos del escenario 1 en el conjunto de datos propio y en el conjunto de datos de Kaggle.

Data Set	Modelos	Precision
Imágenes propias	Resnet50V2	0.92
	DenseNet121	0.89
	InceptionV3	0.94
Imágenes Kaggle	Resnet50V2	0.72
	DenseNet121	0.89
	InceptionV3	0.76

Data Set	Modelos	Precision	Voting
Modelo final Imágenes Propias	Resnet + Inception	0.94587	soft
	Resnet + Inception	0.9252	hard
	DenseNet + Inception	0.9226	hard
	DenseNet + Inception	0.9355	soft
	Resnet + DenseNet	0.91237	hard
	Resnet + DenseNet	0.91752	soft
	Resnet + DenseNet t + Inception	0.9355	hard
	Resnet + DenseNet + Inception	0.9355	soft
	Resnet + Inception	0.8196	hard
Modelo final Imágenes Kaggle	Resnet + Inception	0.7662	soft
	DenseNet t + Inception	0.8816	hard
	DenseNet + Inception	0.8759	soft
	Resnet + DensNet	0.8917	hard
	Resnet + DenseNet	0.8629	soft
	Resnet + DenseNet + Inception	0.82395	hard
	Resnet + DenseNet + Inception	0.8412	soft

En la Tabla 18 se observa que, en el conjunto de imágenes propias, el modelo ResNet+Inception con predicción soft voting obtuvo la mejor precisión (0.94587), por otro lado se observa que la combinación ResNet+DenseNet con predicción hard voting presenta una precisión notable (0.91237) aunque es ligeramente inferior que la anterior. Con respecto a las imágenes de Kaggle se observa que la combinación ResNet+Inception

con predicción soft voting obtiene un accuracy de 0.7662 mientras que la combinación ResNet+DenseNet con hard voting presenta una precisión de 0.8917. En general estos resultados destacan la eficacia de combinar arquitecturas avanzadas como Resnet+DenseNet+Inception en diferentes formas, para optimizar la precisión de las predicciones, siendo para este escenario 1 la combinación ResNet+DenseNet la más adecuada si se quiere lograr una mejor generalización en nuevas imágenes.

Escenario 2 Generación de imágenes al 50%

En la Tabla 19 se observa que, en el conjunto de imágenes propias, el modelo ResNet+DenseNet con predicción soft voting obtuvo la mejor precisión (0.9226), por otro lado se observa que la combinación ResNet+DenseNet+Inception con predicción hard voting presenta una precisión notable (0.9243). Con respecto a las imágenes de Kaggle se observa que la combinación ResNet+DenseNet con predicción soft voting obtiene un accuracy de 0.8802 mientras que la combinación ResNet+DenseNet con hard voting presenta una precisión de 0.8802, para este escenario 2 la combinación ResNet+DenseNet es la más adecuada si se quiere lograr una mejor generalización en nuevas imágenes.

Tabla 19. Métrica de evaluación de los modelos del escenario 2 en el conjunto de datos propio y en el conjunto de datos de Kaggle.

Data Set	Modelos	Precision
Imágenes propias	Resnet50V2	0.93
	DenseNet121	0.88
	InceptionV3	0.90
Imágenes Kaggle	Resnet50V2	0.8196
	DenseNet121	0.8744
	InceptionV3	0.7676

Data Set	Modelos	Precision	Voting
Modelo final Imágenes Propias	Resnet + Inception	0.9226	soft
	Resnet + Inception	0.9209	hard
	DenseNet + Inception	0.8986	hard
	DenseNet + Inception	0.9140	soft
	Resnet + DenseNet	0.9106	hard
	Resnet + DenseNet	0.9226	soft
	Resnet + Densenet + Inception	0.9243	hard
	Resnet + DenseNet + Inception	0.9192	soft
Modelo final Imágenes Kaggle	Resnet + Inception	0.8412	hard
	Resnet + Inception	0.8167	soft
	DenseNet + Inception	0.8643	hard
	DenseNet + Inception	0.8787	soft
	Resnet + DenseNet	0.8715	hard
	Resnet + DenseNet	0.8802	soft
	Resnet + DenseNet + Inception	0.8542	hard
	Resnet + DenseNet + Inception	0.8658	soft

Escenario 3 Generación de imágenes al 300%

En la Tabla 20 se observa que, en el conjunto de imágenes propias, el modelo ResNet+DenseNet con predicción soft voting obtuvo la mejor precisión (0.9626), por otro lado se observa que la combinación ResNet+DenseNet con predicción hard voting presenta una precisión notable (0.9464). Con respecto a las imágenes de Kaggle se observa que la combinación ResNet+DenseNet con predicción soft voting obtiene un accuracy de 0.8961 mientras que la combinación ResNet+DenseNet con hard voting presenta una precisión de 0.8831 para este escenario 3 la combinación ResNet+DenseNet es la más adecuada si se quiere lograr una mejor generalización en nuevas imágenes.

Tabla 20. Métrica de evaluación de los modelos del escenario 3 en el conjunto de datos propio y en el conjunto de datos de Kaggle

Data Set	Modelos	Precision	
Imágenes propias	Resnet50V2	0.95	
	DenseNet121	0.94	
	InceptionV3	0.94	
Imágenes Kaggle	Resnet50V2	0.8037	
	DenseNet121	0.8961	
	InceptionV3	0.7272	

Data Set	Modelos	Precision	Voting
Modelo final Imágenes Propias	Resnet + Inception	0.9516	soft
	Resnet + Inception	0.9587	hard
	DenseNet + Inception	0.9509	hard
	DenseNet + Inception	0.9638	soft
	Resnet + DenseNet	0.9464	hard
	Resnet + DenseNet	0.9626	soft
	Resnet + DenseNet + Inception	0.9638	hard
	Resnet + DenseNet + Inception	0.9658	soft
Modelo final Imágenes Kaggle	Resnet + Inception	0.8095	hard
	Resnet + Inception	0.7705	soft
	DenseNet + Inception	0.8672	hard
	DenseNet + Inception	0.8860	soft
	Resnet + DenseNet	0.8831	hard
	Resnet + DenseNet	0.8961	soft

	Resnet + DenseNet + Inception	0.8311	hard
	Resnet + DenseNet + Inception	0.8441	soft

7.3. Discusión de resultados

El desarrollo tecnológico ha permitido utilizar arquitecturas de aprendizaje profundo para la clasificación de imágenes en los experimentos realizados de esta investigación. Los resultados demostraron la influencia de diversas configuraciones en la precisión y generalización de modelos, considerando tanto conjuntos de datos propios (HNCH) como nuevos (KAGGLE). En seguida, se realizaron las evaluaciones a los modelos propuestos que lograron satisfacer a los objetivos planteados al inicio de la investigación.

En el escenario 1 (sin generación de imágenes), el modelo ResNet+Inception con predicción soft voting mostró un mejor desempeño destacado en el conjunto de datos propio, con una precisión de 0.94587. Esto lo ubica como la combinación más eficiente para datos locales, lo que indica su capacidad para capturar patrones complejos en imágenes citológicas. Sin embargo, en el conjunto de Kaggle, el modelo ResNet+DenseNet con predicción hard voting obtuvo una precisión de 0.8917, lo que sugiere que esta combinación es más robusta para datos nuevos, debido a su habilidad para generalizar en contextos variados.

En el escenario 2 (con generación de imágenes aumentadas al 50%), se observó una mejoría general en la precisión, destacando la combinación ResNet+DenseNet+Inception con hard voting en el conjunto de datos propio con una precisión de 0.9243. Esta mejora está asociada con una mayor cantidad de datos, que optimiza la capacidad del modelo para aprender. En el conjunto de datos de Kaggle, tanto hard como soft voting con

ResNet+DenseNet alcanzaron un accuracy de 0.8802, demostrando que la generación parcial de imágenes mejora la consistencia entre los métodos de votación.

En el escenario 3 (con generación de imágenes al 300%) se alcanzó los mejores resultados, demostrando nuevamente que ResNet+DenseNet con soft voting alcanzó una precisión de 0.9626 en el conjunto de datos propio, y una precisión de 0.8961 en el conjunto de datos de Kaggle. Esto sugiere que el incremento masivo de imágenes mejora la capacidad de generalización, debido a que el modelo experimenta una mayor variedad de ejemplos, lo que reduce el sobreajuste y aumenta la robustez. Los resultados obtenidos en el Escenario 3, en la combinación ResNet+DenseNet con predicción soft voting alcanzó una precisión de 0.9626 en el conjunto de datos propio, mostrando una mejora en comparación con los métodos de diagnóstico computacional reportados por Jiang et al. (2023). Este avance sugiere que el enfoque basado en modelos ensemble puede superar las técnicas tradicionales de aprendizaje profundo, al mejorar la generalización y precisión en este tipo de imágenes.

Los modelos empleados en esta investigación logran precisiones similares al 90% en la clasificación de células benignas y malignas, tal como lo reportaron Elia et al. (2020). Sin embargo, se destaca la aplicación de métodos ensemble como el hard y soft voting, que en el caso del Escenario 1 con datos del conjunto de Kaggle con ResNet+DenseNet y hard voting, alcanzaron una precisión de 0.8917.

En los experimentos de este estudio, la combinación ResNet+DenseNet en el Escenario 3 se acercó mucho a la precisión de 0.96 alcanzada por el modelo U-Net de Aboobacker et al. (2020). Además, el uso del soft voting proporciona una alternativa que iguala la recuperación y especificidad reportadas, como se evidenció en los datos propios. Esto demuestra que los modelos ensemble pueden competir eficazmente con arquitecturas como U-Net en este dominio.

El enfoque basado en redes neuronales convolucionales profundas reportado por Xie et al. (2022) alcanzó una precisión del 91.67% en la clasificación de células malignas y benignas. En esta investigación, el uso de técnicas avanzadas de aprendizaje por

transferencia y combinación de modelos mediante ensemble permitió obtener precisiones superiores, como en el Escenario 3 con imágenes propias, donde ResNet+DenseNet+Inception bajo predicción hard voting logró una precisión de 96.58%. Esto refuerza la relevancia de las técnicas de combinación en mejorar la robustez de los modelos para la clasificación de imágenes.

Los resultados confirman la eficacia de combinar arquitecturas avanzadas como ResNet, DenseNet e Inception. Sin embargo, la dependencia de la generación de imágenes sugiere una posible limitación en términos de recursos computacionales y tiempo de entrenamiento. Además, la variabilidad entre conjuntos propios y nuevos/externos destaca la necesidad de evaluar modelos en contextos diversos para garantizar su aplicabilidad clínica.

Esta investigación aporta evidencia sobre cómo optimizar la precisión de modelos en análisis citológico de líquido pleural mediante la combinación de técnicas de aumento de datos y aprendizaje por transferencia, alineándose con las tendencias actuales en el uso de aprendizaje profundo para aplicaciones médicas.

CAPÍTULO 8. CONCLUSIONES Y TRABAJOS FUTUROS

8.1. Conclusiones

La inteligencia artificial viene transformando los servicios de atención de la salud mejorando la precisión del diagnóstico, permitiendo una detección más temprana de enfermedades y mejorando los resultados.

Los modelos basados en Inteligencia Artificial son de utilidad para diseñar sistemas de apoyo al diagnóstico médico y para la clasificación células cancerosas en imágenes citológicas de líquido pleural.

La creación de un conjunto de datos de imágenes citológicas a partir de muestras de citología de líquido pleural representa una contribución valiosa para la comunidad académica por ser único en el país y al haber sido desarrollado por un equipo multidisciplinario de la Facultad de Ingeniería de Sistemas e Informática de la UNMSM y el Servicio de Anatomía Patológica del HNCH.

La investigación evaluó diferentes arquitecturas de aprendizaje profundo con aprendizaje por transferencia y combinaciones de modelos para la clasificación de imágenes citológicas, destacando la efectividad de las técnicas avanzadas en este campo.

Combinar arquitecturas como ResNet, DenseNet e Inception mediante métodos de votación (hard y soft voting) demostraron ser más precisas que los modelos individuales, sobre todo en escenarios donde se incrementó la cantidad del conjunto de datos mediante generación de imágenes.

Realizar el aumento de datos al 300% fue importante para mejorar la generalización de los modelos, alcanzando precisiones superiores al 96% en el conjunto de datos propio y hasta el 89.61% en el conjunto de datos de Kaggle. Este experimento resaltó la importancia de incluir técnicas de generación de datos en la construcción de modelos robustos para aplicaciones médicas.

Es importante resaltar que los resultados en el conjunto de datos propio fueron consistentemente altos, las combinaciones de modelos también mostraron un buen desempeño en el conjunto de Kaggle, lo que demuestra la capacidad de estos modelos para generalizar en conjunto de datos externos que ayudaría en su aplicación en entornos clínicos reales.

Se demostró la contribución al área de la citología ya que los hallazgos son consistentes con estudios previos y refuerzan el uso de algoritmos de aprendizaje profundo con aprendizaje por transferencia como herramientas valiosas para el diagnóstico en citología. La combinación de modelos avanzados puede ofrecer una precisión comparable o superior a las técnicas tradicionales, con el potencial de integrarse en sistemas automatizados de apoyo a la decisión médica. El modelo de IA ahorrará tiempo, reducirá el agotamiento de los especialistas y mejorará la atención a los pacientes.

Este trabajo demuestra el valor de combinar modelos de aprendizaje profundo, aprendizaje por transferencia y técnicas de aumento de datos para mejorar la precisión y la aplicabilidad clínica en la clasificación de imágenes citológicas de líquido pleural, contribuyendo al avance en el diagnóstico computacional.

- ✓ En relación con el OE1, que consistía en analizar las diferentes técnicas y enfoques utilizados en estudios previos sobre el uso del aprendizaje profundo para clasificar células neoplásicas en imágenes citológicas de líquido pleural, los resultados obtenidos muestran que los modelos ResNet50V2, DenseNet121 e InceptionV3 fueron los seleccionados por su capacidad de clasificación efectiva. Además, se exploraron combinaciones de estos modelos a través de técnicas de votación (hard voting y soft voting) para maximizar la precisión en los conjuntos de datos utilizados (imágenes propias y de Kaggle). Los resultados muestran que la combinación de modelos aporta un mejor rendimiento que los modelos individuales en la mayoría de los casos, lo cual es coherente con estudios previos que indican que el uso de modelos en conjunto puede mejorar la precisión y la generalización del modelo.

- ✓ Respecto al OE2, que tenía como finalidad construir un conjunto de datos locales de imágenes citológicas de líquido pleural, el uso de imágenes propias del Hospital Nacional Cayetano Heredia (HNCH) permitió evaluar la efectividad del modelo en un contexto local. Adicionalmente, se utilizaron imágenes del conjunto de datos de Kaggle para realizar comparaciones y evaluar la generalización. Los resultados reflejan que el conjunto de datos local proporcionó mejores precisiones en los escenarios sin generación de imágenes y con generación al 50%, destacando la importancia de utilizar datos propios que reflejen con mayor precisión la realidad local en los procesos de entrenamiento y evaluación del modelo.
- ✓ En cuanto al OE3, que planteaba el diseño e implementación de un modelo de inteligencia artificial para la clasificación de células neoplásicas, el modelo ResNet+DenseNet con soft voting demostró ser el más efectivo en el escenario con generación de imágenes al 300%, alcanzando una precisión de 0.9626 en imágenes propias. Este resultado sugiere que el uso de técnicas de aumento de datos, como la generación de imágenes, contribuye significativamente a la mejora del rendimiento del modelo. Asimismo, el empleo de combinaciones de modelos con técnicas de votación permite una mejor integración de las fortalezas individuales de cada arquitectura.
- ✓ Finalmente, en el contexto del OE4, la precisión alcanzada en el modelo desarrollado fue 0.9626 con las combinaciones de las técnicas de votación. Luego al someter al conjunto de datos (KAGGLE) a las comparaciones y evaluaciones, se llega a la conclusión que es más generalizable. Esto demuestra que el modelo desarrollado puede considerarse un recurso complementario para el apoyo médico en el diagnóstico citológico, dado que su alto rendimiento podría optimizar los procesos de diagnóstico y reducir la carga laboral de los profesionales de salud.

8.2. Limitaciones

Una de las limitaciones al inicio del trabajo fue que en Perú no existe un conjunto de datos de imágenes citológicas por lo que se decidió a construir un conjunto de datos con

imágenes de un hospital público, presentando algunos inconvenientes administrativos que luego fueron superados.

Otra limitación fue que no se han realizado en Perú modelos IA que clasifiquen células neoplásicas en imágenes citológicas de líquido pleural, por lo que no encontramos investigaciones locales.

8.3. Trabajos futuros

El derrame pleural es una patología que frecuentemente se presenta en los pacientes con cáncer y sin diagnóstico de cáncer por lo que es considerado un problema de salud pública. Existen tareas que se deben promover para ayudar a un diagnóstico efectivo como un trabajo colaborativo entre instituciones públicas, la academia y la empresa privada. Se pudo iniciar desarrollando este tipo de aplicaciones de ayuda al diagnóstico en salud, aquí es donde encontramos nichos para futuras investigaciones como:

Construcción de conjunto de datos de hospitales públicos de todo el país.

Desarrollo e implementación de interfaz que pueda conectar zonas internas del país con los departamentos de Anatomía Patológica de los centros de salud hospitales con infraestructura superior.

Implementar en un hospital público el modelo IA para que se pueda considerar la usabilidad y aceptación de los expertos.

Si bien la sensibilidad de la citología para diagnosticar el derrame pleural maligno es limitada, la evaluación de biomarcadores tumorales en el líquido pleural parece prometedora para mejorar la precisión del diagnóstico, convirtiéndose en una potencial fuente de estudio en el futuro.

8.4. Recomendaciones

Algunas limitaciones y perspectivas futuras están relacionadas al escaso acceso a conjuntos de datos de imágenes citológicas de líquido pleural de acceso público, por lo

que se recomienda trabajar en construir conjuntos de datos locales que ayudará a mejorar la representatividad y aplicabilidad de los diferentes modelos de Inteligencia Artificial. En esta parte, es importante utilizar técnicas de balanceo y transformación de datos para que la generalización del modelo tenga un resultado óptimo.

Futuras investigaciones podrían enfocarse en optimizar los tiempos de entrenamiento y reducir la dependencia de recursos computacionales, además de explorar la integración con otras arquitecturas especializadas como U-Net o Transformers.

CONTRIBUCIONES

Entre las principales contribuciones que destacan en este trabajo de investigación se encuentran:

- ✓ La construcción, desarrollo e implementación de un conjunto de datos locales de muestras de imágenes citológicas de líquido pleural de un hospital público. Este conjunto de datos se podría utilizar en otras investigaciones, en fines educativos para los especialistas que inician su formación en Citología y su actividad profesional.
- ✓ El diseño de un modelo IA para clasificar células neoplásicas en imágenes de citología de líquido pleural.
- ✓ Primer modelo trabajado con un conjunto de datos locales de creación propia de un centro de salud público, que se podrá ampliar su utilidad para diagnóstico citológico de otro tipo de neoplasias y otros temas médicos.

REFERENCIAS

- Aboobacker, S., Vijayasenan, D., Sumam David, S., Suresh, P. K., & Sreeram, S. (2020). A Deep Learning Model for the Automatic Detection of Malignancy in Effusion Cytology. *ICSPCC 2020 - IEEE International Conference on Signal Processing, Communications and Computing, Proceedings*. <https://doi.org/10.1109/ICSPCC50002.2020.9259490>
- Alrafiah, A. R. (2022). Application and performance of artificial intelligence technology in cytopathology. *Acta Histochemica*, 124(4), 151890. <https://doi.org/10.1016/J.ACTHIS.2022.151890>
- Araujo. (2020). *Manejo del paciente con derrame pleural. ¿Cómo estudiarlos para conseguir un diagnóstico específico? (Managing the patient with pleural effusion. How to study them in order to get a specific diagnosis?)*. <https://orcid.org/0000-0002-6559-5370>.
- Aubreville, M., Bertram, C. A., Donovan, T. A., Marzahl, C., Maier, A., & Klopffleisch, R. (2020). A completely annotated whole slide image dataset of canine breast cancer to aid human breast cancer research. *Scientific Data*, 7(1). <https://doi.org/10.1038/s41597-020-00756-z>
- Bashour, S. I., Mankidy, B. J., & Lazarus, D. R. (2022). Update on the diagnosis and management of malignant pleural effusions. In *Respiratory Medicine* (Vol. 196). W.B. Saunders Ltd. <https://doi.org/10.1016/j.rmed.2022.106802>
- Baykal, B., Sarioglu, O. F., & Aydin, O. (2020). The role of cytology in the diagnosis of malignant pleural effusions. *Journal of Cytology*, 37(1), 1–7. https://doi.org/10.4103/JOC.JOC_123_19
- Baykal, E., & Doğan, H. (2017a). *Automated nuclei detection in serous effusion cytology based on machine learning* (M. Felsberg, A. Heyden, & N. Krüger, Eds.; Vol.

- 10425). Springer International Publishing. <https://doi.org/10.1007/978-3-319-64698-5>
- Baykal, E., & Doğan, H. (2017b). *Automated nuclei detection in serous effusion cytology based on machine learning* (M. Felsberg, A. Heyden, & N. Krüger, Eds.; Vol. 10425). Springer International Publishing. <https://doi.org/10.1007/978-3-319-64698-5>
- Bera, K., Schalper, K. A., Rimm, D. L., Velcheti, V., & Madabhushi, A. (2019). Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology. *Nature Reviews Clinical Oncology*, 16(11), 703–715. <https://doi.org/10.1038/s41571-019-0252-y>
- Berkhout, M., Smit, K., & Versendaal, J. (2024). Decision discovery using clinical decision support system decision log data for supporting the nurse decision-making process. *BMC Medical Informatics and Decision Making*, 1, 100. <https://doi.org/https://doi.org/10.1186/s12911-024-02486-3>
- Bhattacharya, S., Reddy Maddikunta, P. K., Pham, Q. V., Gadekallu, T. R., Krishnan S, S. R., Chowdhary, C. L., Alazab, M., & Jalil Piran, M. (2021). Deep learning and medical image processing for coronavirus (COVID-19) pandemic: A survey. *Sustainable Cities and Society*, 65. <https://doi.org/10.1016/j.scs.2020.102589>
- Bielsa, PENADES, EGIDO, & RUE. (2008). *Rentabilidad del estudio citológico del líquido pleural en el derrame maligno*.
- Bishop. (2024). *Foundations and Concepts Deep Learning*.
- Bravo, Brayan, Arteaga, Soto, Iván, Buriticá, & Orozco-Arias. (2021a). Clasificación de cáncer cervical usando redes neuronales convolucionales, transferencia de aprendizaje y aumento de datos. *Revista EIA ISSN 1794-1237 Ed. 35, 18*(2463–0950), 1–12. <https://doi.org/10.24050/reia>

- Bravo, Brayan, H., Arteaga, A., Soto, R. T., Iván, J., Buriticá, P., & Orozco-Arias, S. (2021b). Clasificación de cáncer cervical usando redes neuronales convolucionales, transferencia de aprendizaje y aumento de datos. *Revista EIA ISSN 1794-1237 Ed. 35, 18*(2463–0950), 1–12. <https://doi.org/10.24050/reia>
- Budak, A., Yanarateş, A., & Budak, E. (2023). The contribution of PET/CT to the differentiation of benign and malignant pleural effusion in patients with ovarian carcinoma. *Revista Española de Medicina Nuclear e Imagen Molecular (English Edition)*, 42(6), 353–358. <https://doi.org/10.1016/J.REMNIE.2023.05.001>
- Carmona, M. G.-C. E. P. R. M. (2021). Revisión sistemática sobre la efectividad de la Patología Digital como herramienta de diagnóstico asistencial en anatomía patológica y citopatología. Ministerio de Sanidad. Agencia de Evaluación de Tecnologías Sanitarias. Instituto de Salud Carlos III. *Ministerio de Sanidad. Agencia de Evaluación de Tecnologías Sanitarias. Instituto de Salud Carlos III.*
- Centro Internacional de Investigaciones sobre el Cáncer (CIIC). (2024, February 1). *Crece la carga mundial de cáncer en medio de una creciente necesidad de servicios*. Centro Internacional de Investigaciones Sobre El Cáncer (CIIC).
- Centro Nacional de Epidemiología, P. y C. de E. (2021). *Análisis de situación de salud RRHH del Perú*.
- Chen, Z., Chen, K., Lou, Y., Zhu, J., Mao, W., & Song, Z. (2021). Machine learning applied to near-infrared spectra for clinical pleural effusion classification. *Scientific Reports*, 11(1). <https://doi.org/10.1038/s41598-021-87736-4>
- CMP. (2021). *Informe sobre la distribución de especialistas médicos en el Perú*.
- Dabeer, S., Khan, M. M., & Islam, S. (2019). Cancer diagnosis in histopathological image: CNN based approach. *Informatics in Medicine Unlocked*, 16. <https://doi.org/10.1016/j.imu.2019.100231>

- Deng. (2022). Un estudio de redes neuronales convolucionales para la clasificación de imágenes: modelos y conjuntos de datos. *Actas de La Conferencia Internacional Sobre Big Data, Información y Redes Informáticas* .
- Deng, Dong, Socher, & Li. (2009). *ImageNet: A large-scale hierarchical image database*.
- Dermawan, J. K. T., & Policarpio-Nicolas, M. L. (2020). Malignancies in pleural, peritoneal, and pericardial effusions. *Archives of Pathology and Laboratory Medicine*, 144(9), 1086–1091. <https://doi.org/10.5858/arpa.2019-0429-OA>
- Dixit, R., & Agarwal, K. (2017). Diagnóstico y opciones de tratamiento en los derrames pleurales malignos. *Pulmon de La India*.
- Elia, S., D'Angelo, G., Palmieri, F., Sorge, R., Massoud, R., Cortese, C., Hardavella, G., & De Stefano, A. (2020). A machine learning evolutionary algorithm-based formula to assess tumor markers and predict lung cancer in cytologically negative pleural effusions. *Soft Computing*, 24(10), 7281–7293. <https://doi.org/10.1007/s00500-019-04344-1>
- El-Sheikh, S. A. (2012). The Diagnostic Value of Pleural Fluid Cytology in Benign and Malignant Pleural Effusions. In *Cairo Univ* (Vol. 80, Issue 2). www.medicaljournalofcairouniversity.com
- Ferreiro, L., Toubes, M. E., San José, M. E., Suárez-Antelo, J., Golpe, A., & Valdés, L. (2020). Advances in pleural effusion diagnostics. In *Expert Review of Respiratory Medicine* (Vol. 14, Issue 1, pp. 51–66). Taylor and Francis Ltd. <https://doi.org/10.1080/17476348.2020.1684266>
- Fisher., A. (2014). *Fundamentos del diagnóstico citológico y su base biológica*.
- Gonzales, R., & Wood, R. (2018). *Digital Image Processing* (Pearson Education Limited, Ed.; 4ta ed.).
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*.

- Gottumukkala, M. (2023). Design, Development, and Evaluation of an Automated Solution for Electronic Information Exchange Between Acute and Long-term Postacute Care Facilities: Design Science Research. *JMIR Form Res*, 7, e43758. <https://doi.org/10.2196/43758>
- Gudmundsson, E., Straus, C. M., Li, F., & Armato, S. G. (2020). Deep learning-based segmentation of malignant pleural mesothelioma tumor on computed tomography scans: application to scans demonstrating pleural effusion. *Journal of Medical Imaging*, 7(01), 1. <https://doi.org/10.1117/1.jmi.7.1.012705>
- Guerra, E., & García-Mayor, R. (2018). *RETOS ÉTICOS QUE PLANTEA EL USO DE LA INTELIGENCIA ARTIFICIAL EN EL DIAGNÓSTICO Y TRATAMIENTO CLÍNICO*. *XXIX*, 303–304. <https://doi.org/10.30444/CB.15>
- Gui, C., & Chan, V. (2017). *Machine learning in medicine*.
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition*. <http://arxiv.org/abs/1512.03385>
- Hernández-Sampieri, Fernández, & Baptista. (2014). *Metodología de la investigación*.
- Hevner, A. R., & Chatterjee, S. (2010). *Design Science Research in Information Systems*. 9–22. https://doi.org/10.1007/978-1-4419-5653-8_2
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Q.*, 28(1), 75–105. <https://doi.org/10.2307/25148625>
- Hooper, C., Lee, Y. C. G., & Maskell, N. (2010). Investigation of a unilateral pleural effusion in adults: British Thoracic Society pleural disease guideline 2010. *Thorax*, 65(SUPPL. 2). <https://doi.org/10.1136/thx.2010.136978>
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. *Proceedings - 30th IEEE Conference on Computer Vision*

and *Pattern Recognition, CVPR 2017, 2017-January*, 2261–2269.
<https://doi.org/10.1109/CVPR.2017.243>

Huang, L., Lin, Y., Cao, P., Zou, X., Qin, Q., Lin, Z., Liang, F., & Li, Z. (2024). Automated detection and segmentation of pleural effusion on ultrasound images using an Attention U-net. *Journal of Applied Clinical Medical Physics*, 25(1).
<https://doi.org/10.1002/acm2.14231>

Huang, Yang, & Yang. (2020). Advances in pleural effusion cytology: The application of artificial intelligence. *Journal of Clinical Oncology Research*, 12(4), 145–152.
<https://doi.org/10.1200/JCO.2020.12.4.145>

Huertas, J., Ramírez, M., & Quispe, L. (2022). Limitaciones en el acceso a servicios especializados en oncología en provincias peruanas. *Revista Médica Peruana*, 39(1), 23–30. <https://doi.org/10.17843/rmp.2022.39.1>

INEN. (2022). *Informe anual de cáncer en el Perú*. <https://www.inen.gob.pe>

Instituto Nacional del Cancer. (2020). *Estadísticas del cáncer - NCI*.
<https://www.cancer.gov/espanol/cancer/naturaleza/estadisticas>

Jiang, Zhou, Lin, Chan, Liu, & Chen. (2023). Deep learning for computational cytology: A survey. In *Medical Image Analysis* (Vol. 84). Elsevier B.V.
<https://doi.org/10.1016/j.media.2022.102691>

Kaggle, & Parikshit Sanyal. (2021, November 2). *Body cavity fluid cytology images*. KAGGLE. <https://www.kaggle.com/datasets/cmacus/body-cavity-fluid-cytology-images>

Karadayi, S., & Sahin, E. (2021). Surgical treatment in malignant pleural effusion. *Revista Turca de Cirugía Torácica y Cardiovascular*, 577–585.

- Khin, King, & Kazuhiko. (2017). *K Mean Clustering Based Automated Segmentation of Overlapping Cell Nuclei in Pleural Effusion Cytology Images* (2017th ed.). International Conference on Advanced Technologies for Communications.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. <https://doi.org/10.1145/3065386>
- Kumar, Y., Gupta, S., Singla, R., & Hu, Y. C. (2022). A Systematic Review of Artificial Intelligence Techniques in Cancer Prediction and Diagnosis. In *Archives of Computational Methods in Engineering* (Vol. 29, Issue 4, pp. 2043–2070). Springer Science and Business Media B.V. <https://doi.org/10.1007/s11831-021-09648-w>
- Lecun, Y., Bengio, Y., & Hinton, G. (2015a). Deep learning. In *Nature* (Vol. 521, Issue 7553, pp. 436–444). Nature Publishing Group. <https://doi.org/10.1038/nature14539>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Lecun, Y., Bengio, Y., & Hinton, G. (2015b). Deep learning. In *Nature* (Vol. 521, Issue 7553, pp. 436–444). Nature Publishing Group. <https://doi.org/10.1038/nature14539>
- Li, Y., & Song, P. (2022). Revisión del aprendizaje por transferencia en la clasificación de imágenes médicas. *Revista de Imagen y Gráfica*, 10068961.
- Li, Y., Tian, S., Huang, Y., & Dong, W. (2021). Driverless artificial intelligence framework for the identification of malignant pleural effusion. *Translational Oncology*, 14(1). <https://doi.org/10.1016/j.tranon.2020.100896>
- Light. (2002). *Pleural effusion*. www.nejm.org
- Light, R. W. (2021). *Pleural Diseases* (7th ed.). Wolters Kluwer.

- López-Córdova, Vega-Huerta, Maquen-Niño, Cáceres-Pizarro, Adrianzén-Olano, & Benito-Pacheco. (2025). 2 *IJOE-Construction of a New Data Set of Pleural Fluid Cytological Images for Research*. <https://doi.org/doi.org/10.3991/ijoe.v21i07.54323>
- Ma, S., Zhao, M., Wang, H., Shi, F., Sun, X., Chen, S., & Dai, H. N. (2020). Fused 3-stage image segmentation for pleural effusion cell clusters. *Proceedings - International Conference on Pattern Recognition*, 1934–1941. <https://doi.org/10.1109/ICPR48806.2021.9412567>
- March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, 15(4), 251–266. [https://doi.org/10.1016/0167-9236\(94\)00041-2](https://doi.org/10.1016/0167-9236(94)00041-2)
- Martinez, Alfaro, & Garcia. (2001). Bases Teóricas, Digitalización y Análisis de Imágenes. *ES: Club de Informática Aplicada a La Sociedad Española de Anatomía Patológica, ed., Pamplona*.
- Martinez, & Hernandez. (n.d.). *Desarrollo e implementación de clasificadores de imágenes mediante el empleo de*.
- Martínez, S. M. R., Vidotti, S. A. B. G., & Jorente, M. J. V. (2016). Conceptual representation of digital medical images: Integrating Context and Visual Content | Representación conceptual de imágenes médicas digitales: Integración de contexto y contenido visual. *Revista General de Informacion y Documentacion*, 26(2), 651–672.
- Mdletshe, S., Motshweneng, O. S., Oliveira, M., & Twala, B. (2023). Design science research application in medical radiation science education: A case study on the evaluation of a developed artifact. *Journal of Medical Imaging and Radiation Sciences*, 54(1), 206–214. <https://doi.org/10.1016/j.jmir.2022.11.007>
- Ministerio de Salud, & Centro Nacional de Epidemiología, P. y C. de E. (CDC). (2019). *ANÁLISIS DE SITUACIÓN DE SALUD DEL PERÚ*. www.vitalstrategies.org
- MINSA. (2023a). *SALA SITUACIONAL DE CÁNCER EN EL PERÚ*.

- MINSA. (2023b). *Vigilancia epidemiológica Cáncer Perú 2023*.
- MINSA del Perú. (2021). *Informe anual de recursos humanos en salud*.
<https://www.minsa.gob.pe>
- Montgomery. (2004). *Diseño y análisis de experimentos* (Limusa WILLEY, Ed.; 2da ed.).
- Naciones Unidas. (2020). Influencia de las tecnologías digitales. *Naciones Unidas*.
<https://www.un.org/es/un75/impact-digital-technologies>
- Nguyen, James, I., & Miller, I. (2015). Diagnostic value of tumor antigens in malignant pleural effusion: A meta-analysis. *Christofer, S*.
- Niazi, M. K. K., Parwani, A. V., & Gurcan, M. N. (2019). Digital pathology and artificial intelligence. In *The Lancet Oncology* (Vol. 20, Issue 5, pp. e253–e261). Lancet Publishing Group. [https://doi.org/10.1016/S1470-2045\(19\)30154-8](https://doi.org/10.1016/S1470-2045(19)30154-8)
- Obermeyer, Z., & Lee, T. H. (2017). Lost in Thought — The Limits of the Human Mind and the Future of Medicine. *New England Journal of Medicine*, 377(13), 1209–1211.
<https://doi.org/10.1056/nejmp1705348>
- Oyola, A. (n.d.). *Desigualdad en la distribución de médicos en el Perú*. Editorial Ciencias Médicas. Retrieved June 28, 2024, from
http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-346620210001000003&lng=es&nrm=iso&tlng=es
- Oyonarte, M. (2015). ENFOQUE DIAGNOSTICO EN EL PACIENTE CON DERRAME PLEURAL. *REV. MED. CLIN. CONDES*, 313–324.
- Ozcelik, N., Ozcelik, A. E., Guner Zirih, N. M., Selimoglu, I., & Gumus, A. (2023). Deep learning for diagnosis of malign pleural effusion on computed tomography images. *Clinics*, 78. <https://doi.org/10.1016/j.clinsp.2023.100210>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J.,

- Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. In *The BMJ* (Vol. 372). BMJ Publishing Group. <https://doi.org/10.1136/bmj.n71>
- Park, Chong, Lee, Yim, Seo, Hwang, Kim, Gong, Cho, N. H., Yoo, C. W., & Choi, H. J. (2023). Deep Learning-Based Computational Cytopathologic Diagnosis of Metastatic Breast Carcinoma in Pleural Fluid. *Cells*, 12(14). <https://doi.org/10.3390/cells12141847>
- Park, H. S., Chong, Y., Lee, Y., Yim, K., Seo, K. J., Hwang, G., Kim, D., Gong, G., Cho, N. H., Yoo, C. W., Yoo, C. W., & Choi, H. J. (2023). Deep Learning-Based Computational Cytopathologic Diagnosis of Metastatic Breast Carcinoma in Pleural Fluid. *Cells*, 12(14). <https://doi.org/10.3390/cells12141847>
- Peffer, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 24(3), 45–77. <https://doi.org/10.2753/MIS0742-1222240302>
- Perumal, J., Lee, P., Dev, K., Lim, H. Q., Dinish, U. S., & Olivo, M. (2022). Machine Learning Assisted Real-Time Label-Free SERS Diagnoses of Malignant Pleural Effusion due to Lung Cancer. *Biosensors*, 12(11). <https://doi.org/10.3390/bios12110940>
- Phoulady, H. A., & Mouton, P. R. (2018). *A New Cervical Cytology Dataset for Nucleus Detection and Image Classification (Cervix93) and Methods for Cervical Nucleus Detection*. <http://arxiv.org/abs/1811.09651>
- Porcel, J. M. (2018). Biomarkers in the diagnosis of pleural diseases: a 2018 update. In *Therapeutic Advances in Respiratory Disease* (Vol. 12). SAGE Publications Ltd. <https://doi.org/10.1177/1753466618808660>

- Porcel, J. M., Esquerda, A., Vives, M., & Bielsa, S. (2014). Etiología del derrame pleural: Análisis de más de 3.000 toracocentesis consecutivas. *Archivos de Bronconeumología*, 50(5), 161–165. <https://doi.org/10.1016/j.arbres.2013.11.007>
- Pramesh, C. S., Badwe, R. A., Bhoo-Pathy, N., Booth, C. M., Chinnaswamy, G., Dare, A. J., de Andrade, V. P., Hunter, D. J., Gopal, S., Gospodarowicz, M., Gunasekera, S., Ilbawi, A., Kapambwe, S., Kingham, P., Kutluk, T., Lamichhane, N., Mutebi, M., Orem, J., Parham, G., ... Weiderpass, E. (2022). Priorities for cancer research in low- and middle-income countries: a global perspective. *Nature Medicine*, 28(4), 649–657. <https://doi.org/10.1038/s41591-022-01738-x>
- Pries-Heje, J., & Baskerville, R. (2008). The design theory nexus. *MIS Quarterly: Management Information Systems*, 32(4), 731–755. <https://doi.org/10.2307/25148870>
- Priyank, H., & Ambika, P. (n.d.). *Deep Learning for Cancer Cell Detection and Segmentation: A Survey*.
- Pumplun, L., Peters, F., Gawlitza, J., & Buxmann, P. (2023). Bringing Machine Learning Systems into Clinical Practice: A Design Science Approach to Explainable Machine Learning-Based Clinical Decision Support Systems. *J. Assoc. Inf. Syst.*, 24, 8. <https://api.semanticscholar.org/CorpusID:259870250>
- Recuero, J., Figueroa, S., & Galvez, C. (n.d.). *GUÍA CLÍNICA SECT PARA EL MANEJO DEL DERRAME PLEURAL MALIGNO GRUPO DE TRABAJO DEL DERRAME PLEURAL MALIGNO*.
- Ren, W., Zhu, Y., Wang, Q., Jin, H., Guo, Y., & Lin, D. (2023a). Deep Learning-Based Classification and Targeted Gene Alteration Prediction from Pleural Effusion Cell Block Whole-Slide Images. *Cancers*, 15(3). <https://doi.org/10.3390/cancers15030752>

- Ren, W., Zhu, Y., Wang, Q., Jin, H., Guo, Y., & Lin, D. (2023b). Deep Learning-Based Classification and Targeted Gene Alteration Prediction from Pleural Effusion Cell Block Whole-Slide Images. *Cancers*, 15(3). <https://doi.org/10.3390/cancers15030752>
- Rethlefsen, M. L., Associate Dean, A., Smathers Libraries, G. A., Kirtley, S., Patricia Ayala, A., Moher, D., Page, M. J., & Koffel, J. B. (n.d.). *PRISMA-S: An Extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews*. Running Title: *PRISMA-S Extension*.
- Richardson, W., Wilson, M., Nishikawa, J., & Hayward, R. (1995). *The well-built clinical question: a key to evidence-based decisions*.
- Rivera. (2022). Actualidad en el manejo clínico del derrame pleural maligno. *Oncología (Ecuador)*, 32(1), 100–111. <https://doi.org/10.33821/603>
- Sahn, S. A. (2006). Pleural Effusions of Extravascular Origin. In *Clinics in Chest Medicine* (Vol. 27, Issue 2, pp. 285–308). <https://doi.org/10.1016/j.ccm.2005.12.004>
- Sanyal, P., & Dey, P. (2023). Using a deep learning neural network for the identification of malignant cells in effusion cytology material. *Cytopathology*, 34(5), 466–471. <https://doi.org/https://doi.org/10.1111/cyt.13260>
- Sexauer, R., & Shan, Y. (2022). *Detección, segmentación y clasificación automatizadas de derrame pleural a partir de tomografías computarizadas mediante aprendizaje automático*.
- Shen, K. R., Bribriescio, A., Crabtree, T., Denlinger, C., Eby, J., Eiken, P., Jones, D. R., Keshavjee, S., Maldonado, F., Paul, S., & Kozower, B. (2017). The American Association for Thoracic Surgery consensus guidelines for the management of empyema. *Journal of Thoracic and Cardiovascular Surgery*, 153(6), e129–e146. <https://doi.org/10.1016/j.jtcvs.2017.01.030>

- Shojaei, S., Roy-Chowdhuri, S., Safi, J., & Grosu, H. B. (2021). Cytologic Investigations for the Diagnosis of Malignant Pleural Effusion in Non-small Cell Lung Cancer: State-of-the-art Review for Pulmonologists. *Journal of Bronchology & Interventional Pulmonology*, 28(4). https://journals.lww.com/bronchology/fulltext/2021/10000/cytologic_investigations_for_the_diagnosis_of.12.aspx
- Siegel, R. L., Miller, K. D., Fuchs, H. E., & Jemal, A. (2021). Cancer Statistics, 2021. *CA: A Cancer Journal for Clinicians*, 71(1), 7–33. <https://doi.org/10.3322/caac.21654>
- Silva, R. (2019). Noise Reduction in Cytological Images of Pleural Effusion Samples Using Advanced Filtering Techniques. *IEEE Transactions on Biomedical Engineering*.
- Simon, H. A. (1996). Understanding the Natural and the Artificial Worlds, The Science of Design. *The Sciences of the Artificial*, 1–24, 111–138.
- Simonyan, K., & Zisserman, A. (2014). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. <http://arxiv.org/abs/1409.1556>
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2014). *Going Deeper with Convolutions*. <http://arxiv.org/abs/1409.4842>
- Tetikkurt, C., Yılmaz, N., Tetikkurt, S., Gundogdu, S., & Disci, R. (2018). The value of exfoliative cell cytology in the diagnosis of exudative pleural effusions. *Monaldi Archives for Chest Disease*, 88(3), 10–15. <https://doi.org/10.4081/monaldi.2018.944>
- Tosun, A. B., Yergiyev, O., Kolouri, S., Silverman, J. F., & Rohde, G. K. (2015). Detection of malignant mesothelioma using nuclear structure of mesothelial cells in effusion cytology specimens. *Cytometry Part A*, 87(4), 326–333. <https://doi.org/10.1002/cyto.a.22602>

- Varone, V., & Troncone, G. (2012). Toma de biopsia (cáncer). *Revista de Bioquímica Celular*.
- Vásquez, A., López, R., & Torres, F. (2023). Challenges in implementing artificial intelligence in Peruvian healthcare: A case study in oncology. *Latin American Journal of Medical Informatics*, 15(2), 45–58. <https://doi.org/10.3390/latam-medinfo.v15i2.45>
- Victória, Atkinson, Buschetto, Cerentini, Casimiro, De Miranda, Daltoé, Stemmer, & von Wangenheim, A. (2021). What is the state of the art of computer vision-assisted cytology? A Systematic Literature Review. In *Computerized Medical Imaging and Graphics* (Vol. 91). Elsevier Ltd. <https://doi.org/10.1016/j.compmedimag.2021.101934>
- Vukovic, D., Wang, A., Antico, M., Steffens, M., Ruvinov, I., van Sloun, R. J., Canty, D., Royse, A., Royse, C., Haji, K., Dowling, J., Chetty, G., & Fontanarosa, D. (2023). Automatic deep learning-based pleural effusion segmentation in lung ultrasound images. *BMC Medical Informatics and Decision Making*, 23(1). <https://doi.org/10.1186/s12911-023-02362-6>
- Wang, S., Tan, X., Li, P., Fan, Q., Xia, H., Tian, S., Pan, F., Zhan, N., Yu, R., Zhang, L., Duan, Y., Xu, J., Ma, Y., Chen, W., Li, Y., Zhao, Z., Liu, C., Bao, Q., Yang, L., & Jin, Y. (n.d.). *Differentiation of malignant from benign pleural effusions based on artificial intelligence Pleural disease*. <https://doi.org/10.1136/thoraxjnl-2021-218581>
- Wang, X., Feng, Y., Wang, S., Wang, D., & Cheng, T. C. E. (2025). An explainable lesion detection transformer model for medical imaging diagnosis decision support: Design science research. *Decision Support Systems*, 196, 114492. <https://doi.org/https://doi.org/10.1016/j.dss.2025.114492>
- Wei, T. T., Zhang, J. F., Cheng, Z., Jiang, L., Li, J. Y., & Zhou, L. (2023). Development and validation of a machine learning model for differential diagnosis of malignant

pleural effusion using routine laboratory data. *Therapeutic Advances in Respiratory Disease*, 17. <https://doi.org/10.1177/17534666231208632>

Who. (2023). *World Health Statistics 2023 Monitoring health for the SDGs Sustainable Development Goals HEALTH FOR ALL*. <https://www.who.int/publications/book-orders>.

Widodo, C. E., Adi, K., & Gernowo, R. (2024). A support vector machine approach for identification of pleural effusion. *Heliyon*, 10(1). <https://doi.org/10.1016/j.heliyon.2023.e22778>

Win, Choomchuay, Hamamoto, Raveesunthornkiat, Rangsirattanakul, & Pongsawat. (2018). Computer aided diagnosis system for detection of cancer cells on cytological pleural effusion images. *BioMed Research International*, 2018. <https://doi.org/10.1155/2018/6456724>

Win, K., Choomchuay, S., Hamamoto, K., Okinawa Kagaku Gijutsu Kenkyū Kiban Seibi Kikō, Kokuritsu Oyama Kōgyō Kōtō Senmon Gakkō, IEEE Computer Society, & Institute of Electrical and Electronics Engineers. (2017). *Artificial Neural Network Based Nuclei Segmentation on Cytology Pleural Effusion Images* (K. Win, S. Choomchuay, & K. Hanamoto, Eds.; ICIIBMS 2017). 2017.

Win, K. Y., Choomchuay, S., & Hamamoto, K. (2017). K mean clustering based automated segmentation of overlapping cell nuclei in pleural effusion cytology images. *International Conference on Advanced Technologies for Communications, 2017-Octob*, 265–269. <https://doi.org/10.1109/ATC.2017.8167630>

Xie, X., Fu, C. C., Lv, L., Ye, Q., Yu, Y., Fang, Q., Zhang, L., Hou, L., & Wu, C. (2022). Deep convolutional neural network-based classification of cancer cells on cytological pleural effusion images. *Modern Pathology*, 35(5), 609–614. <https://doi.org/10.1038/s41379-021-00987-4>

- Xue, H. (2021). Tendencias de investigación recientes sobre compresión de modelos y transferencia de conocimiento en CNN. *Conferencia Internacional IEEE 2021 Sobre Informática, Inteligencia Artificial e Ingeniería Electrónica (CSAIEE)*.
- Yadanar, K., Choomchuay, S., IEEE Thailand Section, Electrical Engineering/Electronics, C., & Institute of Electrical and Electronics Engineers. (2017). *Automated Segmentation of Cell Nuclei in Cytology Pleural Fluid Images Using OTSU Thresholding* (IEEE Xplore, Ed.).
- Zeiler, M., & Fergus, R. (2014). “Visualizing and understanding convolutional networks. *Lecture Notes in Computer Science*.
- Zheng, W., & Hu, Z.-D. (2022). Pleural fluid biochemical analysis: the past, present and future. *Clinical Chemistry and Laboratory Medicine (CCLM)*, 61, 921–934. <https://doi.org/10.1515/cclm-2022-0844>
- Zhou, Q., Zhang, X., & Li, W. (2021). Machine learning in pleural effusion diagnosis: A systematic review. *Artificial Intelligence in Medicine*, 124, 102125. <https://doi.org/10.1016/j.artmed.2021.102125>

ANEXOS

ANEXO 1. Equipo multidisciplinario con institución pública.

Señor Doctor,
CAYO VICTOR LEÓN FERNÁNDEZ
 Director de la Unidad de Posgrado (UPG)
 Facultad de Ingeniería de Sistemas e Informática
 Universidad Nacional Mayor de San Marcos
cleonf@unmsm.edu.pe
frida.lopez@unmsm.edu.pe
 Mesa de partes digital
 Teléfono: 995-475-431
Presente.-

Asunto: Inteligencia Artificial aplicada al diagnóstico de Derrame Pleural Maligno

Referencia: Documento s/n, de fecha 15 de julio del 2024. Expediente N°11865/2024

De mi consideración:

Es grato dirigirme a usted para saludarlo y en atención al documento de la referencia, mediante el cual nos invita a participar del proyecto colaborativo de investigación doctoral de la *Mg. Frida Meryda Lopez Cordova*, para desarrollarlo en el servicio de Anatomía Patológica a través de la Unidad de Citología de nuestra institución HNCH.

Al respecto, mucho agradeceré a usted, se sirva presentar el trabajo de tesis titulado: ***"Modelo de Inteligencia Artificial para diagnóstico de células neoplásicas en imágenes citológicas de líquido pleural"*** para ser evaluado por el Comité de Ética del HNCH y puedan aprobar su solicitud; los requisitos a presentar son los siguientes:

1. Solicitud dirigida al Director General por el Investigador Principal solicitando la autorización para la realización del trabajo de investigación, solicitando en el mismo documento la exoneración de pago, por todo concepto si no tiene patrocinador, por revisión.
2. Solicitud dirigida al Presidente del Comité Institucional de Ética en Investigación del HNCH, pidiendo la aprobación del trabajo de investigación.
3. 03 copias del trabajo de Investigación.
4. 03 Consentimientos Informados (de existir).
5. 01 Currículo VITAE del Investigador Principal.
6. **Documento de responsabilidad del trabajo de Investigador, suscrito por el Asesor, quien deberá ser Médico u otro Profesional Nombrado del HNCH, según corresponda.**
7. Anexos: A, C, D, E y F (si el estudio es realizado con drogas), debidamente llenados.
8. Factura de pago por revisión de trabajo de investigación, si cuenta con patrocinador.
9. Documentos presentados en archivador de palanca tamaño oficio y en una memoria USB.

Para mayor información y/o consultas sobre los requisitos contactarse al correo cadl.hch@gmail.com o comunicarse al teléfono 2136000-Anexo 209.

Sin otro particular, hago propicia la oportunidad para expresarle mis sentimientos de mayor consideración y estima.

Atentamente




ANEXO 2. Constancia de revisión y aprobación del proyecto, código 102-2024

	PERÚ Ministerio de Salud	Viceministerio de Prestaciones y Aseguramiento en Salud	Hospital Nacional Cayetano Heredia	
Decenio de la Igualdad de Oportunidades para Mujeres y Hombres "Año del Bicentenario, de la Consolidación de Nuestra Independencia y de la Conmemoración de las Heroicas Batallas de Junín y Ayacucho"				

CONSTANCIA N° 227-2024

EL PRESIDENTE DEL COMITÉ INSTITUCIONAL DE ÉTICA EN LA INVESTIGACIÓN DEL HOSPITAL NACIONAL CAYETANO HEREDIA,

CERTIFICA:

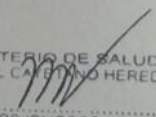
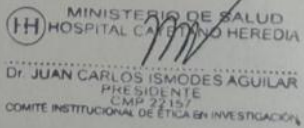
Que el Trabajo de Investigación de la REF: "Modelo de Inteligencia Artificial para diagnóstico de células cancerosas en imágenes citológicas de líquido pleural". Inscrito con Código **102-2024**. Presentado por la Lic. **Frida Mereyda Lopez Cordova** Investigadora Principal ha sido revisado y **APROBADO** por este Comité, en sesión de fecha 12 de diciembre de 2024.

Esta aprobación tendrá vigencia hasta el 11 de diciembre de 2025. Los trámites para su renovación deberán iniciarse por lo menos 30 días antes de su vencimiento.

Participaron en la sesión los siguientes miembros, quienes no presentan conflictos de interés (Según Reglamento del CIEI HNCH año 2022 versión 20 de octubre 2022 - Art. 64, mínimo de participantes para quorum: 05 miembros titulares):

- Dr. Juan Carlos Ismodes Aguilar
- Dr. Javier Renato Puma Medina
- Q.F. Mercedes Jacqueline Portocarrero Reátegui
- Dr. Luis Eduardo Cano Jon
- Psic. Liliana Isabel Cuya Cayo
- Bach. Der. Roberto Hugo Sedano Sedano
- Ing. Ruth Damariz Benito Matias
- Sr. José Leónidas Orcón Bernal

Lima, 12 de diciembre de 2024



Dr. JUAN CARLOS ISMODES AGUILAR
 PRESIDENTE
 CMP 22157
 COMITÉ INSTITUCIONAL DE ÉTICA EN INVESTIGACIÓN

JCIA/kaa


**BICENTENARIO
PERÚ
2024**

www.hospitalcayetano.gob.pe
DIRECCIÓN GENERAL

Av. Honorio Delgado N° 262
 Urb. Ingeniería
 San Martín de Porres
 Lima 31, PERÚ
 Telf.: (511) 754-4990
 Anexo: 477

ANEXO 3. Autorización Institucional del Trabajo de Investigación

	PERÚ Ministerio de Salud	Viceministerio de Prestaciones y Aseguramiento en Salud	Hospital Nacional Cayetano Heredia	
---	------------------------------------	---	------------------------------------	---

Decenio de la Igualdad de Oportunidades para Mujeres y Hombres
 "Año del Bicentenario, de la Consolidación de Nuestra Independencia y de la Conmemoración de las Heroicas Batallas de Junín y Ayacucho"

San Martín de Porres, 12 de diciembre de 2024

CARTA N° 529 - 2024 - DG - 586 - OEGRH- 296 -OADI/HNCH

Señora Licenciada
FRIDA MEREYDA LOPEZ CORDOVA
 Investigadora Principal
 Universidad Nacional Mayor de San Marcos
Presente.-

Asunto : Autorización Institucional del Trabajo de Investigación.

Referencia : Doc. S/N del 20 de setiembre de 2024
 Expediente N° 16330/2024 - HNCH

De mi consideración:

Me dirijo a usted, para saludarla cordialmente y comunicarle que, contando con la aceptación del Departamento de Patología Clínica y Anatomía Patológica y la Aprobación del Comité Institucional de Ética en la Investigación, esta Dirección autoriza la ejecución del Trabajo de Investigación "Modelo de Inteligencia Artificial para diagnóstico de células cancerosas en imágenes citológicas de líquido pleural". Código: **102-2024**.

Esta autorización tendrá vigencia mientras dure el Trabajo; sin embargo, la Constancia de Aprobación deberá ser renovada anualmente por el Comité Institucional de Ética en Investigación (CIEI) del Hospital Cayetano Heredia.

Cabe resaltar que, la Investigadora Principal asumirá la responsabilidad del Trabajo de Investigación y deberá informar al CIEI, de acuerdo a normas vigentes de cualquier enmienda, eventos adversos, avance, cierre y el respectivo informe final, según corresponda.

Es propicia la ocasión, para expresarle los sentimientos de mi especial consideración.

Atentamente,


 J. SILVERA T.


 V°B°


 MINISTERIO DE SALUD
 HOSPITAL NACIONAL CAYETANO HEREDIA
 DR. SEGUNDO CECILIO ACHO MEGO
 DIRECTOR GENERAL
 CMP 027291 RNE 011026

SCAM/JAST/AOPB/kaa
 C.c. DG
 C.c. OEGRH
 C.c. OADI
 C.c. CIEI

 BICENTENARIO PERÚ 2024
 17 DIC

 Con PUNCHE Perú

www.hospitalcayetano.gob.pe DIRECCIÓN GENERAL

Av. Honorio Delgado N° 262
 Urb. Ingeniería
 San Martín de Porres
 Lima 31, PERÚ
 Telf.: (511) 754-4990
 Anexo: 237

ANEXO 4.

Artículo: Revisión Sistemática de la Literatura.

Título: “*Aprendizaje automático para el diagnóstico de células cancerosas en imágenes citológicas de líquido pleural: Una revisión sistemática de la literatura*”

Revista: Revista Peruana de Computación y Sistemas de la FISI - UNMSM.

Estado: Publicado.

REVISTA PERUANA DE COMPUTACIÓN Y SISTEMAS 2024 6(2):79-85
https://doi.org/10.15381/rpcs.v6i2.29281

ISSN Online: 2617-2003
FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

Artículo de Revisión

Aprendizaje automático para el diagnóstico de células cancerosas en imágenes citológicas de líquido pleural: Una revisión sistemática de la literatura

Machine learning for cancer cell diagnosis in pleural fluid cytology images: A systematic literature review

Frida López^{1,a}, Hugo Vega^{1,b}, Gisella Maquén^{2,c}, Ciro Rodríguez^{1,d}, Augusto Bernuy^{1,e}

¹ Universidad Nacional Mayor de San Marcos, Facultad de Ingeniería de Sistemas e Informática, Unidad de Posgrado. Lima, Perú

² Universidad Nacional Pedro Ruiz Gallo, Facultad de Ingeniería de Sistemas e Informática. Lambayeque, Perú

^a Autor de correspondencia: frida.lopez@unmsm.edu.pe, <https://orcid.org/0000-0002-0494-6800>

^b E-mail: hvega@unmsm.edu.pe, <https://orcid.org/0000-0002-4268-5808>

^c E-mail: gmaquenn@unprg.edu.pe, <https://orcid.org/0000-0002-9224-5456>

^d E-mail: cdrodriguez@unmsm.edu.pe, <https://orcid.org/0000-0003-2112-1349>

^e E-mail: abernuy@unmsm.edu.pe, <https://orcid.org/0000-0003-4117-3728>

Abstract

Machine learning is used in medicine to diagnose diseases quickly and accurately, the results of which support the physician in making correct decisions. Pleural effusion, a common disease in which 50% of patients are diagnosed with cancer. The objective was to describe machine learning techniques that are used for the diagnosis of cancer cells in cytological images of pleural fluid. For the systematic review, the PICO strategy and the PRISMA methodology were used. For the research questions, selection criteria were established, identifying 142 articles, selecting 18 articles after filtering. The techniques used were U_Net with 8 articles, Transfer Learning with 4 articles, Support vector machine with 3 articles, CNN with 3 articles, ANN with 3 articles, X-Boost with one article, K-Means with one article and other ML techniques with 4 articles. Regarding the data set, the most used were cytological images in 10 investigations, CT images in 4 investigations, X-ray images in 3 investigations and one investigation with 1 ultrasound. This literature review will support future research to apply the models and techniques, since there are not many studies on the prediction of cancer cells in pleural fluid.

Keywords: Pleural Effusion, Pleural Fluid, Cytological Examination, Machine Learning.

Recibido: 23-10-2024 - Aceptado: 12-12-2024 - Publicado: 30-12-2024

Citar como:

López Frida et al. (2024). Aprendizaje automático para el diagnóstico de células cancerosas en imágenes citológicas de líquido pleural: Una revisión sistemática de la literatura. Revista Peruana de Computación y Sistemas, 6(2):79-85. <https://doi.org/10.15381/rpcs.v6i2.29281>

© Los autores. Este artículo es publicado por la Revista Peruana de Computación y Sistemas de la Facultad de Ingeniería de Sistemas e Informática de la Universidad Nacional Mayor de San Marcos. Este es un artículo de acceso abierto, distribuido bajo los términos de la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0) (<https://creativecommons.org/licenses/by/4.0/deed.es>) que permite el uso, distribución y reproducción en cualquier medio, siempre que la obra original sea debidamente citada de su fuente original.

ANEXO 5.

Artículo: Contribución.

Título: “*Construcción de un Nuevo Conjunto de Datos de Imágenes Citológicas de Líquido Pleural para la Investigación*”,

Revista: Revista Internacional de Ingeniería Biomédica y en Línea (iJOE) – <http://i-joe.org>

Estado: Publicado.

Artículo Q2, eISSN: 2626-8493. Vol. 21 No. 7 (2025)
doi.org/10.3991/ijoe.v21i07.54323 de acceso abierto.

Online-Journals.org


 International Journal of
Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 21 No. 7 (2025) |  OPEN ACCESS
<https://doi.org/10.3991/ijoe.v21i07.54323>

PAPER
**Construction of a New Data Set of Pleural Fluid
Cytological Images for Research**

Frida López-Córdova¹ ,
 Hugo Vega-Huerta¹ ,
 Gisella Luisa Elena
 Maquen-Niño² ,
 Jaime Cáceres-Pizarro³ ,
 Ivan Adrianzén-Olano⁴ ,
 Oscar Benito-Pacheco¹ 

¹Universidad Nacional Mayor de San Marcos, Lima, Perú

²Universidad Nacional Pedro Ruiz Gallo, Lambayeque, Perú

³Hospital Nacional Cayetano Heredia, Lima, Perú

⁴Universidad Nacional Toribio Rodríguez de Mendoza, Chachapoyas, Perú

hvegah@unmsm.edu.pe


 The limited availability of standardized datasets has hindered the implementation of artificial intelligence (AI) models in serous fluid cytology, particularly in pleural fluid analysis. In this paper, we present the construction of a dataset of pleural fluid cytology images. The objective is to generate a dataset of pleural fluid cytologic images validated by two pathologists and classified into five categories for cell diagnosis, which will be used to train AI models. As a methodology, the images represent pleural fluid cytology samples that have been prepared through medical procedures and transferred to slides, providing valuable information when evaluated under the microscope by medical specialists through cytological examination. We documented the entire process for building the pleural fluid cytological image dataset, from image capture, labeling, preprocessing, standardization, and uploading to public platforms. As a result, we obtained a pleural fluid cytology dataset based on the International System (TIS) criteria for reporting serous fluids, classifying samples into AUS, MAL, ND, NFM, and SFM. This dataset is intended to support medical research, deep learning applications in medical image analysis, and improved diagnostic methodologies.

KEYWORDS
 dataset, construction of a dataset, image preprocessing, cytology, cytological images, pleural fluid, cytological diagnosis

ANEXO 6.

Artículo: Contribución.

Título: “*Clasificación de imágenes de citología pleural mediante aprendizaje profundo y votación por conjunto de los modelos ResNet, DenseNet e Inception*”.

Revista: Expert Systems with Applications

Estado: Se envió para un Q1.

Please verify your contribution to Ensemble Deep Learning Models for Classifying Benign and Malignant Cells in Pleural Fluid Cytology Images (Por favor, verifique su contribución a los modelos de aprendizaje profundo de conjuntos para clasificar células benignas y malignas en imágenes de citología de líquido pleural.) Externo Recibidos x

Expert Systems with Applications <em@editorialmanager.com>
para mí ▾

16 sept 2025, 0:35 ☆

inglés (británico) ▾ español ▾ Traducir correo ←

→ Reenviar correo

 Traducido: inglés → español 
Traductor puede cometer errores, así que verifica las traducciones
[Mostrar original](#)

Este es un mensaje automatizado.

Revista: Sistemas expertos con aplicaciones

Título: Modelos de aprendizaje profundo por conjuntos para la clasificación de células benignas y malignas en imágenes citológicas de líquido pleural

Autor de correspondencia: Dr. Hugo Vega-Huerta

Coautores: FRIDA López-Córdova; Gisella Luisa Elena Maquen-Niño; Jaime Cáceres Pizarro; Ciro Rodríguez; David Calderón; Juan Gamarra Moreno; Percy De-la-Cruz-VdV; Lui Grados; Santiago Moquillaza Henríquez

Número de manuscrito: **ESWA-D-25-25010**

