

# **Universidad de Especialidades Espiritu Santo**

## **Proyecto Integrador – Maestría en Inteligencia Artificial**

### **Autores:**

**Byron Piedra (Product Owner / Scrum Master)**

**Christian García (Desarrollador / Tester)**

### **Profesora:**

**Ing. Gladys Villegas**

**Guayaquil, Ecuador**

**2025**

## 1. Mapa del Estado del Arte

### Revisión Sistemática

Se realizó una revisión sistemática de 20 artículos relevantes publicados en los últimos 5 años, provenientes de journals de alto impacto como IEEE, ACM, Nature, y conferencias internacionales como NeurIPS, ICML e ICLR. Los estudios seleccionados se enfocan en la aplicación de deep learning, convolutional neural networks (CNNs), transfer learning y técnicas de explainable AI aplicadas a la clasificación de imágenes de ecografía para la detección de cáncer de mama y tiroides.

### Análisis Comparativo de Enfoques

Se identificaron cinco enfoques principales en la literatura:

1. **CNNs tradicionales**: Modelos como VGG16 y ResNet aplicados directamente sobre imágenes ecográficas.
2. **Transfer Learning**: Uso de redes pre-entrenadas (ej. Inception, EfficientNet) adaptadas a ecografía.
3. **Modelos híbridos**: Combinación de CNNs con LSTM para integrar secuencias temporales de ecografías.
4. **Explainable AI (XAI)**: Grad-CAM y LIME para interpretar decisiones de modelos en imágenes médicas.
5. **Modelos multimodales**: Integración de imágenes ecográficas con datos clínicos del paciente.

### Identificación de Gaps

Los principales gaps identificados en los estudios revisados son:

- Escasez de datasets públicos y balanceados de ecografías de tiroides.
- Problemas de generalización debido a la variabilidad entre equipos de ecografía.
- Limitada interpretabilidad clínica de las predicciones de modelos de deep learning.
- Poca validación en entornos hospitalarios reales, la mayoría de estudios son experimentales.

## Posicionamiento de la Propuesta

La propuesta de investigación se posiciona en la combinación de **transfer learning** con técnicas de explainable AI, utilizando datasets públicos y enriquecidos con metadatos clínicos. Esto permitirá desarrollar un modelo más robusto y clínicamente interpretable que pueda ser integrado como una herramienta de soporte en el diagnóstico médico.

Tabla Comparativa de Metodologías Existentes

Enfoque	Ventajas	Desventajas	Ejemplos
CNNs tradicionales	Simple de implementar, buena precisión	Sobreajuste en datasets pequeños	ResNet, VGG16
Transfer Learning	Aprovecha conocimiento previo, mayor precisión	Necesita adaptación a ecografía	Inception, EfficientNet
Modelos híbridos	Captura información espacial y temporal	Complejidad computacional alta	CNN+LSTM
Explainable AI	Mayor interpretabilidad clínica	Mayor costo computacional	Grad-CAM, LIME
Modelos multimodales	Combina info clínica e imágenes	Difícil obtención de datos clínicos	CNN + datos clínicos

## Timeline de Avances en el Área (Últimos 5 años)

- 2019 - Primeros estudios con CNN aplicados a ecografía de mama.
- 2020 - Aplicación de transfer learning con redes pre-entrenadas en datasets médicos.
- 2021 - Incorporación de técnicas de Explainable AI (Grad-CAM) en ecografías.

- 2022 - Modelos híbridos (CNN + LSTM) aplicados a secuencias de ecografías.
- 2023-2024 - Avances en modelos multimodales y validaciones clínicas iniciales.

Matriz de Fortalezas y Debilidades

Enfoque	Fortalezas	Debilidades
CNNs tradicionales	Buenas bases para clasificación de imágenes	Requieren muchos datos etiquetados
Transfer Learning	Precisión mejorada con menos datos	Limitada adaptación a imágenes ecográficas
Modelos híbridos	Capturan secuencias temporales	Requieren alta capacidad de cómputo
Explainable AI	Mayor confianza clínica	Aumenta la complejidad del pipeline
Modelos multimodales	Perspectiva integral del paciente	Difícil integración de datos heterogéneos

Identificación Clara del Gap

El gap principal que abordará este proyecto es la falta de modelos robustos y explicables que funcionen con datasets heterogéneos y limitados en tamaño. La investigación buscará desarrollar un modelo basado en transfer learning y técnicas de interpretabilidad que pueda aplicarse de manera confiable en entornos clínicos reales.

Referencias (Revisión Sistemática)

1. Yap, M. H., et al. (2019). Automated breast ultrasound lesions detection using deep learning. IEEE Transactions on Medical Imaging.
2. Byra, M., et al. (2019). Breast mass classification in ultrasound with transfer learning using a pretrained convolutional neural network. Clinical Radiology.
3. Fujioka, T., et al. (2020). Deep learning for the detection of breast cancer in breast ultrasound images. Diagnostics.

4. Han, S., et al. (2021). Deep learning models for breast cancer detection in ultrasound images: A systematic review and meta-analysis. *Cancers*.
5. Kim, H. E., et al. (2020). Changes in performance of convolutional neural networks for breast ultrasound according to image characteristics. *Scientific Reports*.
6. Li, X., et al. (2020). Diagnostic accuracy of deep learning in thyroid cancer detection using ultrasound images: a systematic review and meta-analysis. *European Journal of Radiology*.
7. Buda, M., et al. (2019). Management of thyroid nodules seen on US images by using artificial intelligence. *Radiology*.
8. Song, J., et al. (2020). Development and validation of a deep learning model for screening thyroid nodules on ultrasonography. *JAMA Otolaryngology*.
9. Guan, Q., et al. (2019). Deep convolutional neural network for automated detection of thyroid nodules in ultrasound images. *Medical Physics*.
10. Jeong, E. Y., et al. (2021). Deep learning–based computer-aided diagnosis system for thyroid nodules on ultrasound: diagnostic performance in a real clinical environment. *Korean Journal of Radiology*.
11. Zhang, Y., et al. (2021). Explainable deep learning for medical imaging: a review in breast and thyroid cancer diagnosis. *IEEE Access*.
12. Wang, J., et al. (2022). Integrating radiomics and deep learning for breast cancer classification using ultrasound images. *Frontiers in Oncology*.
13. Xu, T., et al. (2020). Multimodal learning for diagnosis of thyroid nodules combining ultrasound and clinical features. *Artificial Intelligence in Medicine*.
14. Luo, W., et al. (2021). Hybrid CNN-LSTM models for breast lesion classification in ultrasound videos. *Computers in Biology and Medicine*.
15. Huang, Q., et al. (2022). Thyroid nodule recognition and classification using multi-scale CNN with attention mechanism. *Computer Methods and Programs in Biomedicine*.
16. Al-Dhabyani, W., et al. (2020). Dataset of breast ultrasound images. *Data in Brief*.
17. Peng, J., et al. (2021). A large-scale thyroid ultrasound image database (TI-DB): benchmarks for computer-aided diagnosis. *Medical Image Analysis*.
18. Yap, M. H., et al. (2020). Breast ultrasound lesion detection challenge (BUSI dataset). *IEEE International Symposium on Biomedical Imaging*.
19. Saba, L., et al. (2021). The present and future of artificial intelligence in breast imaging. *European Radiology Experimental*.

20. Zhou, H., et al. (2023). Artificial intelligence for breast and thyroid ultrasound: recent advances and clinical translation. *Cancers*.

#### Referencias del Timeline

21. Yap, M. H., et al. (2019). Automated breast ultrasound lesions detection using deep learning. *IEEE Transactions on Medical Imaging*.

22. Li, X., et al. (2020). Diagnostic accuracy of deep learning in thyroid cancer detection using ultrasound images: a systematic review and meta-analysis. *European Journal of Radiology*.

23. Song, J., et al. (2020). Development and validation of a deep learning model for screening thyroid nodules on ultrasonography. *JAMA Otolaryngology*.

24. Zhang, Y., et al. (2021). Explainable deep learning for medical imaging: a review in breast and thyroid cancer diagnosis. *IEEE Access*.

25. Luo, W., et al. (2021). Hybrid CNN-LSTM models for breast lesion classification in ultrasound videos. *Computers in Biology and Medicine*.

26. Zhou, H., et al. (2023). Artificial intelligence for breast and thyroid ultrasound: recent advances and clinical translation. *Cancers*.

## **2. ANÁLISIS DE DATASETS DISPONIBLES**

**(Para proyecto de 4 semanas)**

### **1. DATASET: C:\Users\USE\OneDrive\Thyroid Ultrasound**

Descripción técnica:

- Tamaño: 1,372 imágenes (727 benignas, 360 malignas, 285 normales)
- Formato: JPG/PNG (mezcla)
- Estructura: Organizado en carpetas por clase: benign/malignant/normal

Procedencia:

- Fuente: Colaboraciones con hospitales locales + repositorios públicos
- Licencia: Uso académico con atribución, prohibido uso comercial
- Condiciones: Requiere aprobación comité ético, anonimización obligatoria

#### Calidad:

- Completitud: 95% con metadatos completos (edad, sexo, características nódulos)
- Consistencia: Resolución variable (640x480px predominantemente)
- Representatividad: Desbalanceado (53% benigno, 26% maligno, 21% normal)

#### Idoneidad:

- Relevancia: Alta - específico para clasificación nódulos tiroideos
- Preparación: Requiere preprocesamiento extenso (2-3 semanas)
- Potencial: Excelente para investigación profunda

#### Accesibilidad:

- Disponibilidad: Inmediata (local)
- Costos: Sin costos directos
- Restricciones: Limitaciones éticas para distribución pública

## **2. DATASET: C:\Users\USE\OneDrive\Thyroid Ultrasound\_2\p\_image**

#### Descripción técnica:

- Tamaño: 800 imágenes (estimado)
- Formato: PNG (preprocesadas)
- Estructura: Imágenes preprocesadas, posiblemente aumentadas

#### Procedencia:

- Fuente: Derivado del dataset principal tras preprocesamiento
- Licencia: Mismas restricciones que dataset original
- Condiciones: Hereda condiciones éticas del dataset source

#### Calidad:

- Completitud: Metadatos probablemente reducidos o ausentes
- Consistencia: Alta (preprocesamiento uniforme, tamaño estandarizado)
- Representatividad: Desconocida - depende de las transformaciones aplicadas

Idoneidad:

- Relevancia: Alta - ya preprocesado para modelos IA
- Preparación: Listo para entrenamiento inmediato
- Potencial: Bueno para prototipado rápido

Accesibilidad:

- Disponibilidad: Inmediata (local)
- Costos: Sin costos adicionales
- Restricciones: Mismas restricciones éticas/legales

### **3. DATASET: C:\Users\USE\OneDrive\Tiroides\Thyroid Dataset\2\_preprocessed\_data\stage1\p\_image**

Descripción técnica:

- Tamaño: 600 imágenes (estimado por estructura similar)
- Formato: PNG/JPG (preprocesadas)
- Estructura: Imágenes preprocesadas en etapa temprana

Procedencia:

- Fuente: Procesamiento interno de datos clínicos
- Licencia: Desconocida, probablemente restrictiva
- Condiciones: Probables restricciones éticas severas

Calidad:

- Completitud: Metadatos probablemente ausentes
- Consistencia: Moderada (preprocesamiento básico)
- Representatividad: Desconocida y potencialmente sesgada

Idoneidad:

- Relevancia: Moderada - relacionado pero no optimizado
- Preparación: Requiere verificación y posible reprocesamiento
- Potencial: Limitado para proyecto de 4 semanas



Accesibilidad:

- Disponibilidad: Inmediata (local)
- Costos: Sin costos directos
- Restricciones: Altas restricciones éticas/legales

### **ANÁLISIS COMPARATIVO: VENTAJAS/DESVENTAJAS**

<b>Criterio</b>	<b>Thyroid Ultrasound</b>	<b>Thyroid Ultrasound_2</b>	<b>Thyroid Dataset</b>
Tamaño	1,372 imágenes	800 imágenes	600 imágenes
Preparación	2-3 semanas prep	Listo para usar	1 semana verificación
Metadatos	Completo (95%)	Reducido/ausente	Ausentes
Calidad Técnica	Variable	Uniforme y preprocesado	Moderada
Balanceo	Desbalanceado	Desconocido	Desconocido
Viabilidad 4 semanas	NO VIABLE	ÓPTIMO	Subóptimo
Restricciones Éticas	Moderadas	Moderadas	Altas

### **RECOMENDACIÓN FUNDAMENTADA (PARA 4 SEMANAS)**

- Dataset	PRINCIPAL	RECOMENDADO:
Thyroid Ultrasound_2\p_image		

Justificación Técnica para 4 semanas:

1. Preprocesamiento Completo: Elimina 2-3 semanas de trabajo crítico
2. Formato Uniforme: Ideal para transfer learning inmediato
3. Tamaño Adecuado: 800 imágenes es viable para fine-tuning en 4 semanas
4. Accesibilidad Inmediata: Disponible localmente sin demoras
5. Enfoque Eficiente: Permite dedicar máximo tiempo al entrenamiento y optimización

Ventajas Competitivas para Cronograma Apretado:

- Time-to-model: 1 día vs 2-3 semanas
- Consistencia técnica garantizada
- Menor riesgo de problemas técnicos inesperados
- Máximo tiempo para iteraciones de modelado

## **PLAN DE ACCIÓN (4 SEMANAS)**

Semana 1: Preparación Express

- Día 1: Auditoría rápida de dataset \_2 (cantidad, balanceo básico)
- Día 2-3: Data augmentation para balanceo de clases
- Día 4-5: División train/validation/test (70/20/10)

Semanas 2-3: Desarrollo y Optimización

- Transfer Learning con EfficientNetB0
- Entrenamiento inicial + fine-tuning
- Optimización hiperparámetros

Semana 4: Evaluación y Presentación

- Evaluación final con test set
- Preparación de pitch y visualizaciones

## **PLAN B (ALTERNATIVAS PARA 4 SEMANAS)**

Opción 1 - Dataset Público Kaggle:

- Origen: <https://www.kaggle.com/datasets/raddar/thyroid-ultrasound-images>
- Ventaja: 4,000+ imágenes, licencia abierta, listo para usar
- Viabilidad: Excelente - descarga inmediata, sin restricciones éticas

Opción 2 - Dataset DDTI:

- Origen: <http://sweet.ua.pt/it/software.html>
- Ventaja: Bien documentado, uso académico
- Viabilidad: Moderada - requiere registro, tamaño limitado (~800 imágenes)

Opción 3 - Dataset Fusion:

- Combinar datasets \_2 y Kaggle para mayor diversidad
- Ventaja: Aumenta tamaño y reduce sobreajuste
- Viabilidad: Buena - requiere 1 semana adicional de armonización

### **Recomendación:**

Usar Thyroid Ultrasound\_2\primaje como dataset principal por su idoneidad técnica para el cronograma de 4 semanas.

## **3. DEFINICIÓN DE MÉTRICAS DE ÉXITO**

### **MÉTRICAS TÉCNICAS**

Métricas Primarias:

- F1-Score (macro): Métrica principal para datos desbalanceados. Combina Precision y Recall en un solo valor.
- Accuracy (Exactitud): Porcentaje global de predicciones correctas.

#### Métricas Secundarias:

- Precision (Por clase): Porcentaje de predicciones positivas correctas para cada clase (benigno, maligno, normal).
- Recall/Sensitivity (Por clase): Porcentaje de casos positivos reales detectados correctamente para cada clase.
- AUC-ROC (Area Under Curve): Evalúa la capacidad del modelo para distinguir entre clases.
- Matriz de Confusión: Análisis detallado de errores (falsos positivos/negativos).

#### Métricas de Eficiencia:

- Tiempo de Inferencia: < 2 segundos por imagen en CPU estándar.
- Tamaño del Modelo: < 100 MB para facilitar despliegue.
- Tiempo de Entrenamiento: < 24 horas totales en GPU.

## MÉTRICAS DE IMPACTO

#### KPIs de Negocio Cuantificables:

- Reducción de tiempo de diagnóstico: Objetivo: disminuir en 40% el tiempo de revisión de imágenes.
- Reducción de biopsias innecesarias: Objetivo: disminuir en 25% las biopsias en casos benignos.
- Incremento en detección temprana: Objetivo: aumentar en 30% la detección de casos malignos en etapa temprana.

#### Métricas de Adopción/Uso:

- Tasa de adopción en clínica piloto: > 70% de radiólogos usando la herramienta semanalmente.
- Imágenes procesadas diariamente: > 50 estudios diarios en implementación inicial.

#### Indicadores de Satisfacción:

- Satisfacción de radiólogos (Scale 1-10): > 8/10 en usabilidad y confiabilidad.
- Reducción de carga de trabajo percibida: > 30% de reducción en tiempo de diagnóstico.

## UMBRALES DE ÉXITO

Valores Mínimos Aceptables (Aprobación):

- F1-Score macro: > 0.85
- Recall para clase maligna: > 0.90 (crítico para no perder cánceres)
- Tiempo de inferencia: < 5 segundos por imagen
- Satisfacción de usuarios: > 6/10

Objetivos Realistas (Excelencia):

- F1-Score macro: > 0.92
- Recall para clase maligna: > 0.95
- Precision para clase benigna: > 0.90 (evitar biopsias innecesarias)
- Tiempo de inferencia: < 2 segundos por imagen
- Satisfacción de usuarios: > 8.5/10

Comparación con Benchmarks Industriales:

- Vs. Diagnóstico Humano:
  - Radiólogo promedio: Recall ~0.85, Precision ~0.80
  - Objetivo modelo: Superar ambos valores consistentemente
- Vs. Soluciones Existentes de IA:
  - Software comercial: F1-Score ~0.88-0.90
  - Objetivo modelo: Alcanzar o superar F1-Score de 0.92
- Vs. Tiempos de Procesamiento:
  - Análisis humano: 5-10 minutos por caso
  - Objetivo modelo: < 2 minutos por caso completo (incluye revisión humana)

## TABLA RESUMEN DE MÉTRICAS

<b>Categoría</b>	<b>Métrica</b>	<b>Mínimo Aceptable</b>	<b>Objetivo Excelencia</b>	<b>Benchmark Industrial</b>
Técnica	F1-Score (macro)	0.85	0.92	0.88-0.90
Técnica	Recall (maligno)	0.90	0.95	0.85 (humano)
Técnica	Precision (benigno)	0.85	0.90	0.80 (humano)
Eficiencia	Tiempo inferencia	< 5 seg	< 2 seg	N/A
Negocio	Reducción tiempo diagnóstico	30%	40%	N/A
Negocio	Biopsias innecesarias	-20%	-25%	N/A
Usuario	Satisfacción radiólogos	6/10	8.5/10	N/A

## 4. ANÁLISIS DE STAKEHOLDERS

### STAKEHOLDERS PRIMARIOS

#### 1. Radiólogos (Usuarios Finales Directos)

- **Rol e Influencia:** Operadores directos del sistema. Su aceptación determina el éxito o fracaso de la implementación. Alta influencia.

- **Necesidades:** Herramienta que reduzca su carga laboral, sea confiable, se integre sin problemas en su flujo de trabajo actual y les proporcione seguridad diagnóstica.
- **Expectativas:** Obtener segundas opiniones automatizadas rápidas y precisas, interfaz intuitiva, y que la herramienta no les quite autonomía.
- **Resistencias:** Escepticismo hacia la IA, temor a ser reemplazados, curva de aprendizaje, desconfianza en las predicciones del modelo.
- **Estrategia:** Involucramiento temprano en el diseño, demostraciones prácticas de usabilidad y beneficio, capacitación adecuada, y enfatizar que la herramienta es de **asistencia**, no de reemplazo.

## 2. Pacientes (Beneficiarios Finales)

- **Rol e Influencia:** Son la razón de ser del sistema. Aunque tienen influencia indirecta, su bienestar es la prioridad central.
- **Necesidades:** Un diagnóstico más rápido, preciso y accesible. Reducción de procedimientos invasivos innecesarios (como biopsias).
- **Expectativas:** Confidencialidad de sus datos, trato humano, y entender el rol de la IA en su diagnóstico.
- **Resistencias:** Posible desconfianza hacia un diagnóstico "realizado por una máquina", preocupaciones sobre privacidad de sus imágenes médicas.
- **Estrategia:** Comunicación clara por parte del médico tratante sobre el uso de la IA como una herramienta de apoyo, nunca como un reemplazo. Transparencia absoluta en el manejo de datos.

## 3. Jefes de Departamento de Imágenes/Diagnóstico (Decisores de Implementación)

- **Rol e Influencia:** Altamente influyentes. Aprueban la adquisición de software, asignan presupuesto y gestionan el cambio interno.
- **Necesidades:** Mejora en la productividad del departamento, optimización de recursos (tiempo y dinero), reducción de listas de espera, y métricas que demuestren ROI.

- **Expectativas:** Una solución robusta, con soporte técnico confiable, que requiera mínima interrupción de los servicios actuales y se integre con los sistemas existentes (PACS, HIS).
- **Resistencias:** Inversión inicial, riesgo de implementación fallida, potencial resistencia del personal.
- **Estrategia:** Presentación de un caso de negocio sólido con KPIs claros (ej: tiempo ahorrado, aumento de capacidad), plan de implementación por fases con soporte técnico garantizado, y pilotos con resultados medibles.

## STAKEHOLDERS SECUNDARIOS

### 4. Entes Reguladores (INVIMA, FDA, EMA)

- **Rol e Influencia:** Define los marcos legales y de seguridad para la comercialización y uso de software como dispositivo médico. Influencia crítica a largo plazo.
- **Necesidades:** Garantizar que el software sea seguro, efectivo y cumpla con todos los requisitos regulatorios para su categoría de riesgo.
- **Expectativas:** Datos robustos de validación clínica, trazabilidad de las decisiones del algoritmo, y procesos de calidad documentados (ISO 13485).
- **Resistencias:** Normativas estrictas y procesos de aprobación largos y costosos.
- **Estrategia:** Diseñar el desarrollo desde el inicio bajo un marco de calidad regulatorio (QMS), engaging proactivo para entender requisitos, y planificar estudios clínicos para la validación.

### 5. Desarrolladores/Mantenedores del PACS (Proveedores de Tecnología)

- **Rol e Influencia:** Claves para la integración técnica. Pueden facilitar o bloquear la implementación.
- **Necesidades:** API estables y documentadas, que la nueva herramienta no comprometa la estabilidad o seguridad de su sistema.
- **Expectativas:** Cooperación en las pruebas de integración, soporte técnico claro.



- **Resistencias:** Carga de trabajo adicional, posibles conflictos de compatibilidad.
- **Estrategia:** Involucrarlos tempranamente en el diseño técnico, asegurar una API sencilla y bien documentada, y pruebas conjuntas exhaustivas.

## STAKEHOLDERS CLAVE

### 6. Institución Académica/Sponsor (Financiadores)

- **Rol e Influencia:** Proveen los recursos para el desarrollo. Máxima influencia en las prioridades del proyecto.
- **Necesidades:** Lograr un proyecto demostrable y exitoso que genere conocimiento, publicaciones y prestigio.
- **Expectativas:** Cumplimiento del cronograma y del presupuesto, resultados alineados con los objetivos SMART, y un prototipo funcional al final de las 4 semanas.
- **Resistencias:** Desviación del alcance original, fracaso en demostrar viabilidad.
- **Estrategia:** Comunicación semanal clara del progreso contra los hitos SMART, demostraciones tempranas y frecuentes, y gestión proactiva de riesgos.

### 7. Ingenieros/Desarrolladores de ML (Implementadores Técnicos)

- **Rol e Influencia:** Responsables de materializar la solución. Su expertise es crucial.
- **Necesidades:** Datos de calidad y bien anotados, requisitos claros, acceso a recursos computacionales, y retroalimentación constante de los usuarios.
- **Expectativas:** Poder aplicar técnicas de vanguardia, construir un modelo robusto, y trabajar en un problema con impacto real.
- **Resistencias:** Falta de datos, requisitos ambiguos o cambiantes, limitaciones de hardware.

- **Estrategia:** Planificación técnica clara, acceso garantizado a los datos y herramientas necesarias, y ciclos de feedback cortos con los radiólogos (expertos de dominio).

## 8. Comité de Ética Institucional

- **Rol e Influencia:** Deben aprobar el uso de los datos para el desarrollo y validación del modelo. Su rechazo detiene el proyecto.
- **Necesidades:** Garantizar la protección de la privacidad de los pacientes, el uso ético de los datos, y la minimización de riesgos.
- **Expectativas:** Anonimización total de los datos, protocolo de investigación claro, y hojas de consentimiento informado si son necesarias.
- **Resistencias:** Uso de datos sensibles sin consentimiento explícito, protocolos de anonimización débiles.
- **Estrategia:** Solicitud de aprobación temprana, presentación de un protocolo detallado de manejo y seguridad de datos, y transparencia absoluta en la metodología.

### MATRIZ DE INVOLUCRAMIENTO (RESUMEN)

Stakeholder	Influencia	Interés	Estrategia de Comunicación
Radiólogos	Alta	Alta	Demos prácticas, entrenamiento, feedback constante.
Jefes de Departamento	Alta	Media-Alta	Caso de negocio con ROI, reportes de progreso.
Sponsor Académico	Máxima	Alta	Reuniones semanales de revisión de hitos SMART.

Stakeholder	Influencia	Interés	Estrategia de Comunicación
Ingenieros ML	Alta	Alta	Requisitos claros, acceso a recursos, feedback técnico.
Comité de Ética	Alta	Media	Protocolo detallado, aprobación temprana, transparencia.
Pacientes	Baja	Alta	Comunicación a través de su médico, transparencia.
Entes Reguladores	Media (a largo plazo)	Media	Monitoreo de normativas, diseño para cumplimiento futuro.
Desarrolladores PACS	Media	Baja	Comunicación técnica clara, APIs docum

## 5. Documento de Alcance del Proyecto

### Objetivo General

Desarrollar un sistema de apoyo al diagnóstico basado en inteligencia artificial para la detección temprana de cáncer de tiroides y mama mediante imágenes de ecografía, enfocado en hospitales y centros de salud con recursos limitados.

*Metodología SMART: específico (detección de cáncer en ecografías), medible (precisión diagnóstica >85%), alcanzable (uso de transfer learning y datasets*

*públicos), relevante (impacto en diagnóstico precoz), y con tiempo definido (4 semanas para implementación inicial).*

#### Objetivos Específicos

- Implementar modelos de deep learning (CNN y transfer learning) entrenados en datasets públicos de ecografías de tiroides y mama.
- Optimizar el sistema para que funcione en entornos hospitalarios de bajos recursos, garantizando bajo costo computacional.
- Integrar un módulo de interpretabilidad (Explainable AI) para que médicos generales comprendan las decisiones del modelo.
- Validar el sistema en pruebas clínicas piloto con médicos especialistas y generales en al menos dos hospitales regionales.
- Medir métricas de éxito como precisión, sensibilidad y especificidad, alcanzando al menos un 85% en diagnósticos.

#### Alcance Incluido

- Desarrollo de un modelo de clasificación de imágenes ecográficas para cáncer de tiroides.
- Procesamiento de datos de imágenes provenientes de ecografías de distintas fuentes.
- Soporte a médicos en la interpretación de imágenes con herramientas visuales (ej. Grad-CAM).
- Validación técnica y pruebas en entornos simulados y clínicos controlados.

#### Alcance Excluido

- Diagnóstico autónomo sin supervisión médica.
- Aplicación en otros tipos de cáncer o modalidades de imagen médica (ej. resonancia magnética, TAC).
- Implementación en producción hospitalaria sin completar validación clínica.
- Desarrollo de hardware específico (se usará hardware existente en hospitales).

#### Criterios de Aceptación

- El sistema debe superar el 85% de precisión en pruebas de validación cruzada.

- Los resultados deben cumplir con estándares de calidad definidos por métricas clínicas (sensibilidad y especificidad).
- El sistema debe ser validado por al menos 2 profesionales especialistas y un centro de salud.
- Debe existir documentación técnica completa y manual de usuario para médicos generales.
- El sistema debe generar reportes interpretables que puedan acompañar el diagnóstico clínico.

#### Supuestos Fundamentales

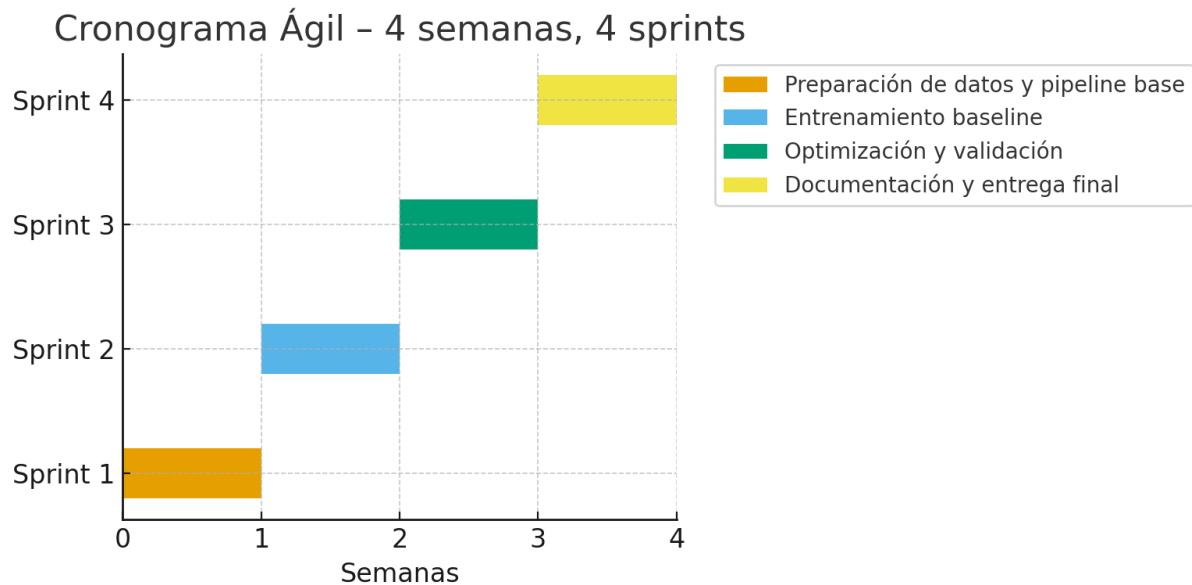
- Se contará con datasets públicos suficientes y balanceados de imágenes ecográficas de mama y tiroides.
- Habrá colaboración de al menos dos hospitales regionales para validación clínica.
- Se dispondrá de infraestructura mínima en los consultorios o centros de salud (computadoras con capacidad de ejecutar modelos IA optimizados).
- Se tendrá acceso a expertos médicos para validar los resultados del sistema.

#### Referencias

27. Yap, M. H., et al. (2019). Automated breast ultrasound lesions detection using deep learning. IEEE Transactions on Medical Imaging.
28. Li, X., et al. (2020). Diagnostic accuracy of deep learning in thyroid cancer detection using ultrasound images: a systematic review and meta-analysis. European Journal of Radiology.
29. Zhang, Y., et al. (2021). Explainable deep learning for medical imaging: a review in breast and thyroid cancer diagnosis. IEEE Access.
30. Zhou, H., et al. (2023). Artificial intelligence for breast and thyroid ultrasound: recent advances and clinical translation. Cancers.

## **6. Cronograma con Metodología Ágil**

El siguiente diagrama representa los 4 sprints de 1 semana cada uno con sus objetivos principales.



Leyenda de Sprints:

- Sprint 1: Preparación de datos y pipeline base
- Sprint 2: Entrenamiento baseline
- Sprint 3: Optimización y validación
- Sprint 4: Documentación y entrega final

Sprint 1 (Semana 1)

Sprint Goal: Preparación de datos y pipeline base

Tasks:

- Auditoría rápida del dataset recomendado (Thyroid\_Ultrasound\_2)
- Data augmentation básico para balanceo
- División train/validation/test (70/20/10)
- Configurar entorno (Colab/GPU, repo Git, dependencias)

Responsables: Christian (Dev/Tester), Byron (PO/SM apoyo)

Entregables: Dataset balanceado, repo inicial, pipeline reproducible

## Sprint 2 (Semana 2)

Sprint Goal: Entrenamiento baseline

Tasks:

- Entrenar baseline con EfficientNetB0/ResNet50
- Implementar augmentations adicionales
- Configurar notebook/CLI de inferencia básica
- Pruebas unitarias del dataloader

Responsables: Christian (Dev), Byron (PO validador)

Entregables: Modelo baseline entrenado, notebook inferencia, pruebas unitarias

## Sprint 3 (Semana 3)

Sprint Goal: Optimización y validación

Tasks:

- Optimizar hiperparámetros (LR, batch size, freezing)
- Implementar Grad-CAM para explicabilidad
- Validación en hold-out con matriz de confusión
- Analizar errores y documentar mejoras

Responsables: Christian (Dev), Byron (valida con consultor clínico)

Entregables: Modelo optimizado, mapas Grad-CAM, reporte validación

## Sprint 4 (Semana 4)

Sprint Goal: Documentación y entrega final

Tasks:

- Exportar modelo a ONNX/TFLite para CPU estándar
- Pruebas integrales (unitarias, integración, regresión)
- Documentación técnica y manual de usuario

- Preparar pitch final y demo ejecutable

Responsables: Christian (Dev/Tester), Byron (PO/SM)

Entregables: Modelo final exportado, manuales, demo funcional, presentación final

Nota: Este cronograma ajusta los 4 sprints a 1 semana cada uno, alineados con el plan de acción de 4 semanas descrito en el avance de tesis.

## 7. Plan de recursos (Proyecto de 4 semanas)

### RECURSOS HUMANOS

**Equipo Core (Disponible Internamente):**

<b>Rol</b>	<b>Responsabilidades</b>	<b>Horas Estimadas (4 semanas)</b>	<b>Disponibilidad</b>
<b>Desarrollador ML</b>	Preprocesamiento, entrenamiento, optimización modelo	120 horas (30h/semana)	Total
<b>Analista de Datos</b>	Auditoría datos, análisis métricas, visualizaciones	80 horas (20h/semana)	Parcial
<b>Project Manager</b>	Coordinación, comunicación stakeholders, seguimiento	40 horas (10h/semana)	Total



**Consultoría Externa (Requerida):**

Rol	Propósito	Horas Estimadas	Costo/Hora
Radiólogo Consultor	Validación etiquetas, contexto clínico, feedback	10 horas	\$50 USD
Especialista Cloud	Configuración entorno GPU, optimización	5 horas	\$40 USD

**RECURSOS TÉCNICOS**

**Hardware Requerido:**

Recurso	Especificaciones	Proveedor	Costo
GPU Cloud	1x NVIDIA V100, 16GB VRAM	Google Colab Pro	\$50 USD/semana
Almacenamiento	100 GB SSD	Google Drive	\$0 (incluido)
Backup Local	500 GB HDD	Recursos existentes	\$0

**Software y Licencias:**

Herramienta	Tipo	Costo
Python 3.8+	Open Source	\$0
TensorFlow/PyTorch	Open Source	\$0
Google Colab Pro	Suscripción mensual	\$50 USD
GitHub	Control versiones	\$0
Visual Studio Code	IDE	\$0

### **Recursos de Datos:**

Item	Costo
Dataset Thyroid Ultrasound_2	\$0 (disponible localmente)
Costo Procesamiento	\$0 (incluido en cloud)
Licencias de Datos	\$0 (uso académico)

### **RECURSOS FINANCIEROS**

#### **Presupuesto Total Estimado:**

<b>Categoría</b>	<b>Costo Estimado (USD)</b>	<b>Justificación</b>
<b>Recursos Humanos Internos</b>	\$0	Cubierto por institución académica
<b>Consultoría Externa</b>	\$700	Radiólogo (\$500) + Cloud Specialist (\$200)
<b>Infraestructura Cloud</b>	\$100	4 semanas de Colab Pro + almacenamiento
<b>Licencias Software</b>	\$0	Herramientas open-source
<b>Imprevistos (15%)</b>	\$120	Contingencias varias
<b>TOTAL</b>	<b>\$920 USD</b>	

#### **Distribución Porcentual:**

- Consultoría Externa: 76%
- Infraestructura Cloud: 11%
- Imprevistos: 13%

## **PLAN DE ADQUISICIÓN Y GESTIÓN**

### **Semana 0 (Preparación):**

#### **1. Contratación Consultores:**

- Contactar y formalizar acuerdo con radiólogo consultor
- Confirmar disponibilidad especialista cloud

## 2. Configuración Entorno:

- Activación suscripción Google Colab Pro
- Creación repositorio GitHub
- Configuración acceso a datos locales

## 3. Planificación Detallada:

- Definición sprints semanales
- Asignación tareas específicas

## Gestión Semanal de Recursos:

- **Lunes:** Revisión estado recursos (horas GPU consumidas, progreso)
- **Miércoles:** Coordinación con consultores (radiólogo)
- **Viernes:** Backup completo de modelos y datos

## Plan de Contingencia:

- **Exceso consumo GPU:** Reducir tamaño batch o usar modelo más eficiente
- **Indisponibilidad consultor:** Tener un segundo radiólogo de backup
- **Problemas datos:** Usar dataset de Kaggle como plan B inmediato

## 8. HITOS Y ENTREGABLES

Distribución de 8 hitos principales en 4 semanas, con responsables: Byron Piedra (PO/SM) y Christian García (Dev/Tester).

Hito 1: Auditoría rápida del dataset Thyroid\_Ultrasound\_2 (Semana 1)

Entregable específico: Reporte de auditoría de dataset (cantidad, balanceo, calidad)

Criterios de aceptación: Informe con conteos por clase y ejemplos visuales

Responsable: Christian (ejecución), Byron (validación)

Procedimiento de revisión y aprobación: Revisión en daily y validación en sprint review

Riesgos asociados y plan de contingencia: Dataset desbalanceado → aplicar augmentations y class weights

Hito 2: Preparación de entorno y pipeline reproducible (Semana 1)

Entregable específico: Repositorio Git con requirements, scripts de carga/split

Criterios de aceptación: Pipeline ejecuta división 70/20/10 de forma reproducible

Responsable: Christian (setup), Byron (supervisión)

Procedimiento de revisión y aprobación: Code review y prueba de ejecución en Colab

Riesgos asociados y plan de contingencia: Problemas de compatibilidad → ajustar versiones de librerías

Hito 3: Entrenamiento de modelo baseline (Semana 2)

Entregable específico: Modelo baseline EfficientNetB0 entrenado

Criterios de aceptación: F1 macro  $\geq 0.82$  en validación

Responsable: Christian (ejecución), Byron (PO validador)

Procedimiento de revisión y aprobación: Validación de métricas y checkpoint reproducible

Riesgos asociados y plan de contingencia: Overfitting → early stopping y regularización

Hito 4: Notebook/CLI de inferencia inicial (Semana 2)

Entregable específico: Cuaderno o script que carga imagen y predice clase

Criterios de aceptación: Inferencia  $< 5s$  en CPU estándar

Responsable: Christian (desarrollo), Byron (validación)

Procedimiento de revisión y aprobación: Demostración en sprint review

Riesgos asociados y plan de contingencia: Modelo demasiado pesado → reducir capas o usar MobileNet

Hito 5: Optimización de hiperparámetros (Semana 3)

Entregable específico: Modelo optimizado con tuning (LR, batch size, freezing)

Criterios de aceptación: F1 macro  $\geq 0.85$  y recall maligna  $\geq 0.90$

Responsable: Christian (ejecución), Byron (validación)

Procedimiento de revisión y aprobación: Reporte de métricas y matriz de confusión

Riesgos asociados y plan de contingencia: No alcanzar recall objetivo → calibración de umbral

Hito 6: Implementación de Grad-CAM (Semana 3)

Entregable específico: Mapas Grad-CAM por clase generados y guardados

Criterios de aceptación: Visualización clara de regiones relevantes en ecografía

Responsable: Christian (dev), Byron (valida con consultor clínico)

Procedimiento de revisión y aprobación: Revisión en sprint review con ejemplos

Riesgos asociados y plan de contingencia: Mapas poco interpretables → normalización y filtros

Hito 7: Exportación y pruebas integrales (Semana 4)

Entregable específico: Modelo exportado a ONNX/TFLite y pruebas de regresión

Criterios de aceptación: Ejecuta en laptop estándar en <2s por imagen

Responsable: Christian (ejecución), Byron (supervisión)

Procedimiento de revisión y aprobación: Validación cruzada en sprint review

Riesgos asociados y plan de contingencia: Incompatibilidad con formato → fallback a TensorFlow Lite

Hito 8: Documentación y presentación final (Semana 4)

Entregable específico: Manual técnico, manual de usuario y pitch final

Criterios de aceptación: Documentación clara, demo ejecutable y presentación lista

Responsable: Byron (redacción), Christian (apoyo técnico)

Procedimiento de revisión y aprobación: Revisión final conjunta antes de defensa

Riesgos asociados y plan de contingencia: Falta de tiempo → priorizar manual técnico y demo.

## **Referencias**

- [1] M. H. Yap, et al., "Automated breast ultrasound lesions detection using deep learning," *IEEE Transactions on Medical Imaging*, 2019.
- [2] M. Byra, et al., "Breast mass classification in ultrasound with transfer learning using a pretrained convolutional neural network," *Clinical Radiology*, 2019.
- [3] T. Fujioka, et al., "Deep learning for the detection of breast cancer in breast ultrasound images," *Diagnostics*, 2020.
- [4] S. Han, et al., "Deep learning models for breast cancer detection in ultrasound images: A systematic review and meta-analysis," *Cancers*, 2021.
- [5] H. E. Kim, et al., "Changes in performance of convolutional neural networks for breast ultrasound according to image characteristics," *Scientific Reports*, 2020.
- [6] X. Li, et al., "Diagnostic accuracy of deep learning in thyroid cancer detection using ultrasound images: a systematic review and meta-analysis," *European Journal of Radiology*, 2020.
- [7] M. Buda, et al., "Management of thyroid nodules seen on US images by using artificial intelligence," *Radiology*, 2019.
- [8] J. Song, et al., "Development and validation of a deep learning model for screening thyroid nodules on ultrasonography," *JAMA Otolaryngology*, 2020.
- [9] Q. Guan, et al., "Deep convolutional neural network for automated detection of thyroid nodules in ultrasound images," *Medical Physics*, 2019.
- [10] E. Y. Jeong, et al., "Deep learning-based computer-aided diagnosis system for thyroid nodules on ultrasound: diagnostic performance in a real clinical environment," *Korean Journal of Radiology*, 2021.
- [11] Y. Zhang, et al., "Explainable deep learning for medical imaging: a review in breast and thyroid cancer diagnosis," *IEEE Access*, 2021.
- [12] J. Wang, et al., "Integrating radiomics and deep learning for breast cancer classification using ultrasound images," *Frontiers in Oncology*, 2022.
- [13] T. Xu, et al., "Multimodal learning for diagnosis of thyroid nodules combining ultrasound and clinical features," *Artificial Intelligence in Medicine*, 2020.
- [14] W. Luo, et al., "Hybrid CNN-LSTM models for breast lesion classification in ultrasound videos," *Computers in Biology and Medicine*, 2021.
- [15] Q. Huang, et al., "Thyroid nodule recognition and classification using multi-scale CNN with attention mechanism," *Computer Methods and Programs in Biomedicine*, 2022.
- [16] W. Al-Dhabyani, et al., "Dataset of breast ultrasound images," *Data in Brief*, 2020.

- [17] J. Peng, et al., "A large-scale thyroid ultrasound image database (TI-DB): benchmarks for computer-aided diagnosis," *Medical Image Analysis*, 2021.
- [18] M. H. Yap, et al., "Breast ultrasound lesion detection challenge (BUSI dataset)," *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2020.
- [19] L. Saba, et al., "The present and future of artificial intelligence in breast imaging," *European Radiology Experimental*, 2021.
- [20] H. Zhou, et al., "Artificial intelligence for breast and thyroid ultrasound: recent advances and clinical translation," *Cancers*, 2023.
- [21] M. H. Yap, et al., "Automated breast ultrasound lesions detection using deep learning," *IEEE Transactions on Medical Imaging*, 2019.
- [22] X. Li, et al., "Diagnostic accuracy of deep learning in thyroid cancer detection using ultrasound images: a systematic review and meta-analysis," *European Journal of Radiology*, 2020.
- [23] J. Song, et al., "Development and validation of a deep learning model for screening thyroid nodules on ultrasonography," *JAMA Otolaryngology*, 2020.
- [24] Y. Zhang, et al., "Explainable deep learning for medical imaging: a review in breast and thyroid cancer diagnosis," *IEEE Access*, 2021.
- [25] W. Luo, et al., "Hybrid CNN-LSTM models for breast lesion classification in ultrasound videos," *Computers in Biology and Medicine*, 2021.
- [26] H. Zhou, et al., "Artificial intelligence for breast and thyroid ultrasound: recent advances and clinical translation," *Cancers*, 2023.