

The background features several large, overlapping geometric shapes. In the top-left corner, there is a maroon triangle pointing towards the center. In the top-right corner, there is a larger maroon triangle pointing towards the center. At the bottom, there are three overlapping triangles: a grey one on the left, a maroon one in the center, and a grey one on the right, all pointing upwards.

DIAGNÓSTICO AVANZADO DE IMÁGENES **APLICANDO TÉCNICAS DE INTELIGENCIA** **ARTIFICIAL**

Nombre de la materia

Proyecto Integrador en Inteligencia Artificial

Nombre del alumno

Christian Garcia y Byron Piedra

Nombre de la tarea

Proyecto Integrador

Nombre del Profesor

Ing. Gladys Villegas

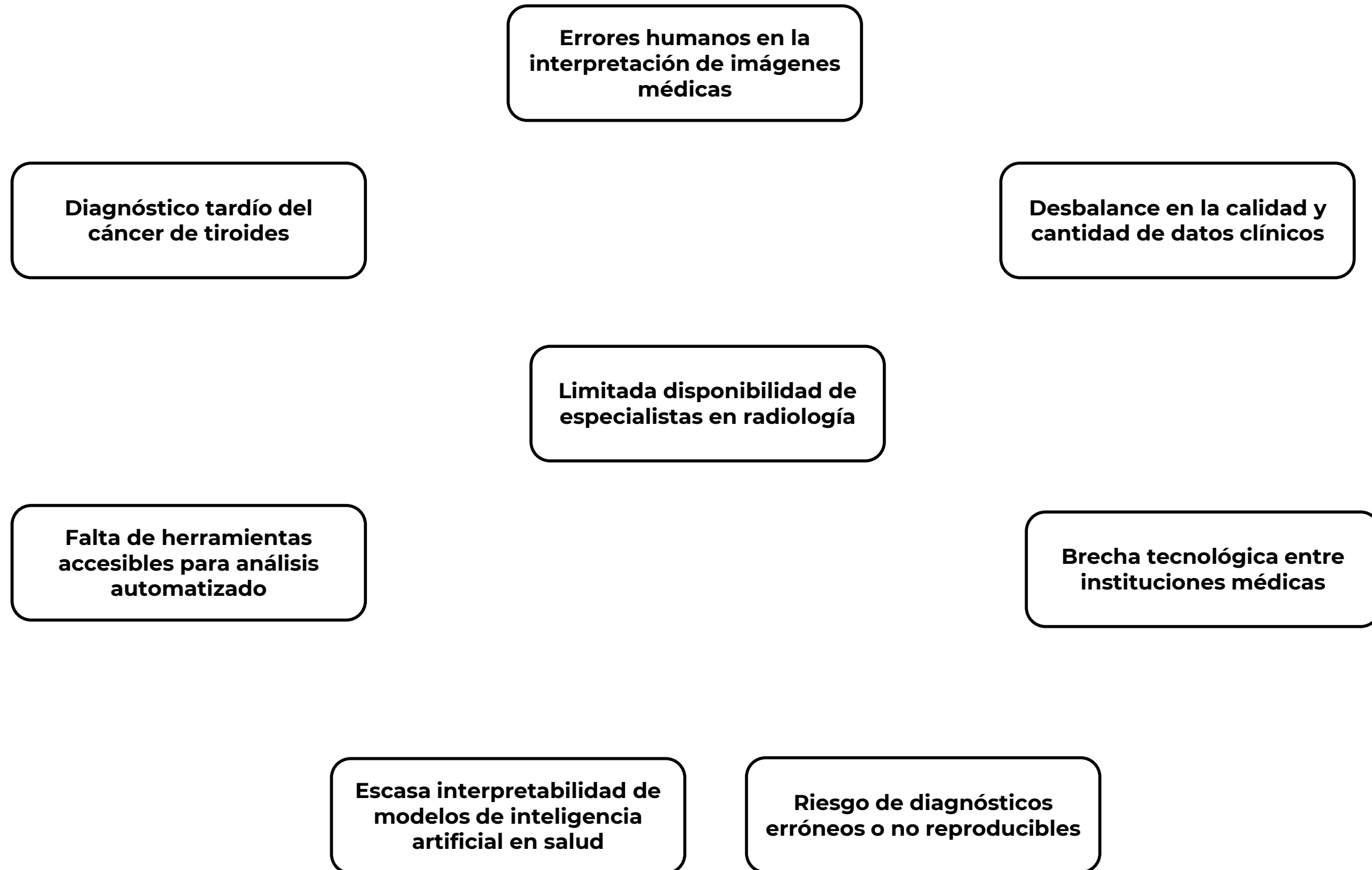
Fecha

11/12/2025

Introducción

El proyecto desarrolla un sistema de apoyo al diagnóstico médico basado en inteligencia artificial para la detección temprana del cáncer de tiroides mediante ecografía. Se implementó un flujo completo que abarca preprocesamiento de imágenes, extracción de características, modelado y evaluación. El modelo final, un Random Forest optimizado, alcanzó un 64% de precisión y superó al baseline en todas las métricas clave. El sistema utiliza datos anónimos y públicos, asegurando criterios éticos y de privacidad.

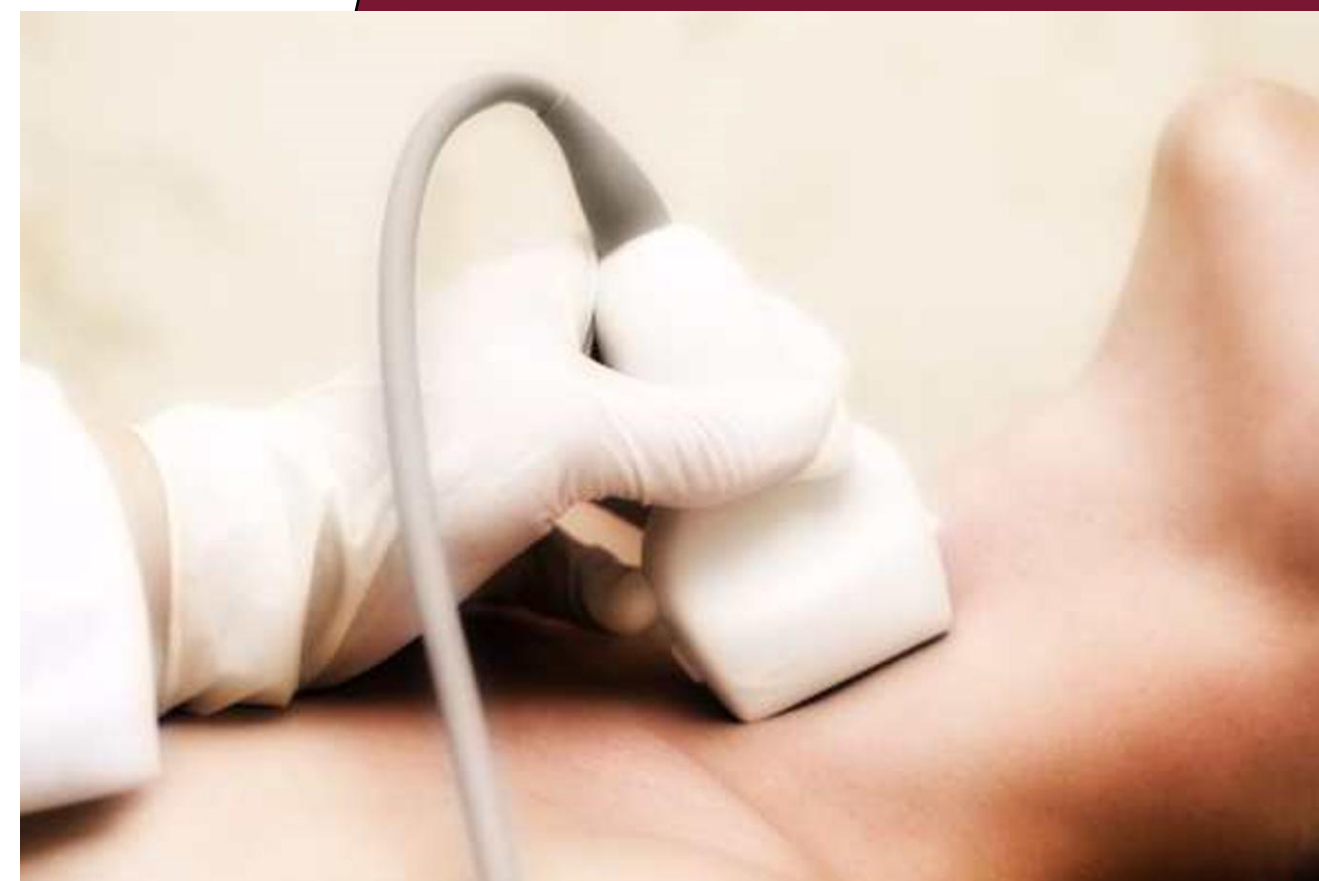
PROBLEMÁTICA



»

Motivación y relevancia

La motivación principal de este proyecto surge de la necesidad de mejorar los procesos de diagnóstico temprano del cáncer de tiroides, una enfermedad cuya detección oportuna incrementa significativamente las probabilidades de éxito en el tratamiento. En numerosos centros médicos, especialmente en regiones con recursos limitados, la falta de especialistas en radiología y la alta carga de trabajo generan retrasos diagnósticos y variabilidad en la interpretación de las ecografías. Esta realidad plantea la urgencia de integrar herramientas tecnológicas que apoyen la labor médica sin reemplazarla.



Datos y Metodología

Dataset Utilizado

- El proyecto empleó un dataset compuesto por 637 imágenes ecográficas tiroideas.
- Clasificadas en dos categorías: benignas y malignas
- Split estratificado (70% entrenamiento, 20% validación/ajustes y 10% prueba.
- Las imágenes para el entrenamiento se sacaron de la siguiente dirección drive:

/content/drive/MyDrive/p_1_image

Preprocesamiento Realizado

- Carga y limpieza de imágenes
- Redimensionamiento (224×224)
- Normalización (0–1)
- Conversión a RGB
- División Train/Val/Test
- Preparación de batches

Datos y Metodología

FLUJO DE PROCESAMIENTO PRINCIPAL:

1. **Carga** (Cargar_dataset_completo_avanzado)
2. **Preprocesamiento** (Cargar y preprocesar_imagen_avanzado)
3. **Extracción características** (Extraer características avanzadas_completas)
4. **Análisis EDA**
5. **Entrenamiento modelo** (Crear modelo de predicción compatible)
6. **Predicción** (Predecir Imágenes tiroides)
7. **Diagnóstico** (Generar diagnostico detallado)



Modelo

1. Random Forest (Optimizado)
2. Clasificación binaria: *benign vs malignant*
3. Buen desempeño con datasets pequeños
4. Robusto al ruido
5. Reduce sobreajuste

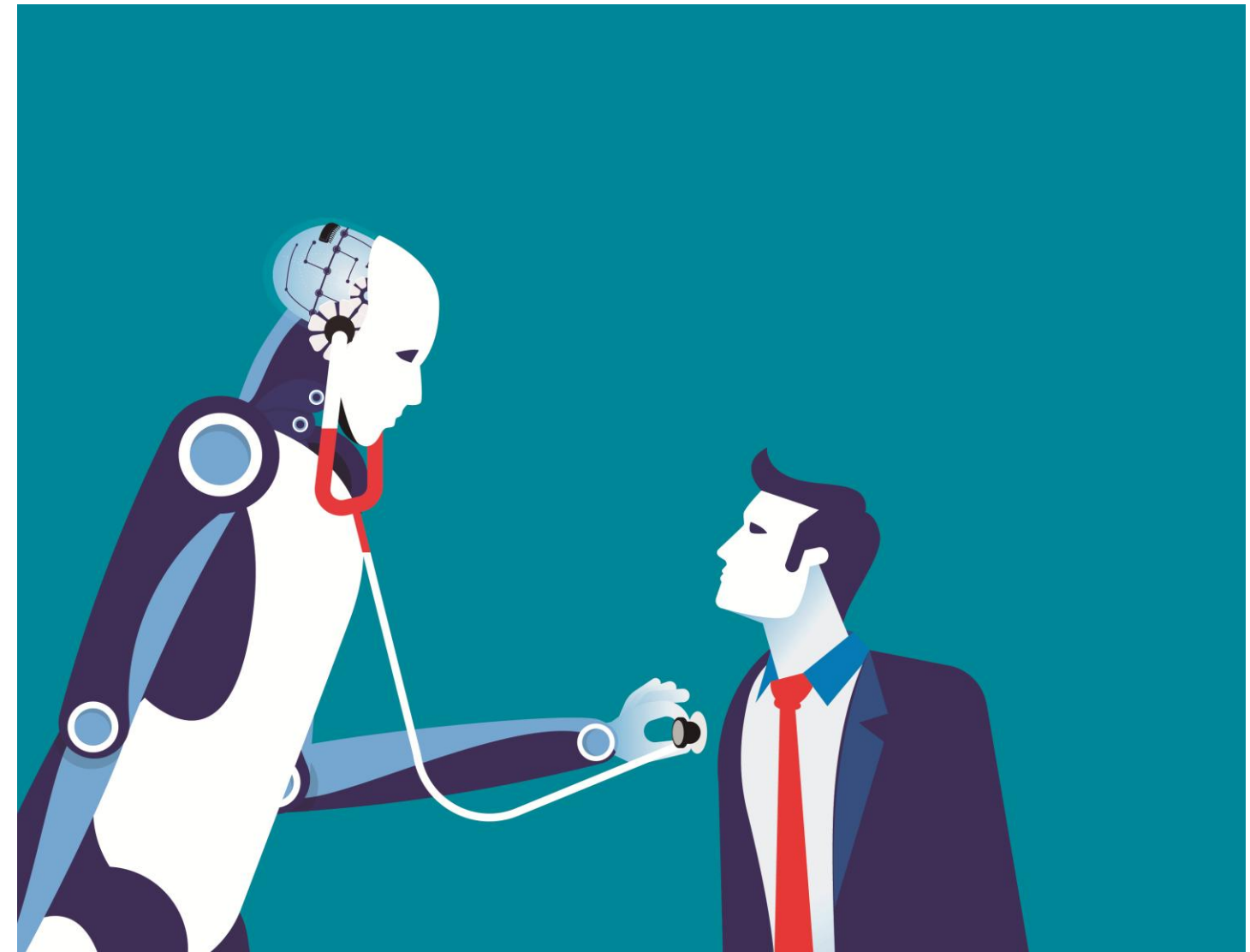
Arquitectura

1. Ensamble de árboles de decisión
2. Votación mayoritaria entre árboles
3. Uso de subconjuntos aleatorios de datos y atributos
4. Optimización de hiperparámetros (*n_estimators*, *max_depth*)
5. Estabilidad y mejor generalización

Datos y Metodología

Justificación de Decisiones Técnicas

1. Selección de RF Optimizado
2. CNN Optimizado con EfficientnetB0
3. Redimensionamiento/Normalización
0-1/Conversión RGB
4. Ajustes de hiperparámetros mediante
GridSearch y RandomSearch
5. Validación cruzada
6. Optimización por accuracy y precision



DEMO

Inicio de Sesión

Usuario

admuser

Contraseña

admuser

Ingresar

Login

EDA

Entrenamiento

Home

Predicción por Imagen

Panel de Anlisis Tiroideo con IA

Elige en la Barra Izquierda la opcion deseada

localhost:8501/EDA

← → ↺ localhost:8501/EDA

🔍 traductor de google... 🖨 Generated Image N... 🗣 ChatGPT 📺 TradingView 📺 Mapa de calor de ac... 📊 RED Ning Invierta P... 📈 MacroTrends Incom... 🌐 Yah

Login

EDA

Entrenamiento

Home

Predicción por Imagen

🔍

Análisis EDA

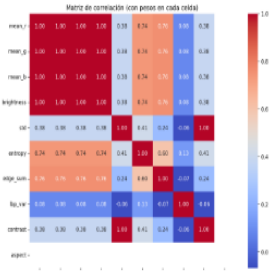
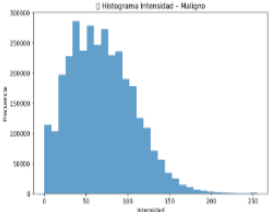
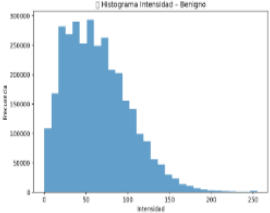
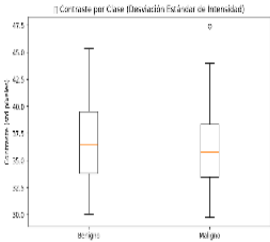
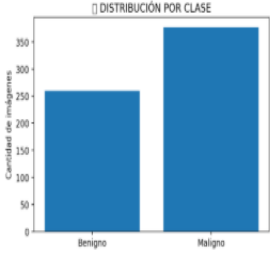
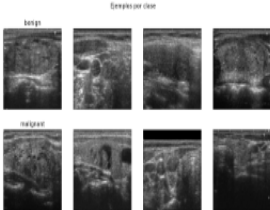
Pulsa para ejecutar el EDA. Se mostrarán gráficos y logs capturados.

Ejecutar EDA

Análisis EDA

Resumen general de la muestra de datos y tipo de datos.

Resumen de la muestra



Ver logs de consola (EDA)

Ver logs de consola (EDA)

```
Total imágenes: 637
label
malignant    377
benign       260
Name: count, dtype: int64
Extrayendo features para EDA: 0% |

Estadísticas de features:
count      mean      std      min      25%
mean_r    637.0    6.797486e+01    1.269129e+01    35.400969    5.942492e+01    6.7376
mean_g    637.0    6.797486e+01    1.269129e+01    35.400969    5.942492e+01    6.7376
mean_b    637.0    6.797486e+01    1.269129e+01    35.400969    5.942492e+01    6.7376
brightness 637.0    2.665681e-01    4.976975e-02    0.138827    2.330389e-01    2.6415
std        637.0    1.455547e-01    1.698747e-02    0.095115    1.333319e-01    1.4481
entropy    637.0    6.731508e+00    2.035605e-01    5.506553    6.613694e+00    6.7535
edge_sum   637.0    1.657958e+06    4.132191e+05    468690.000000    1.362210e+06    1.6393
lbp_var    637.0    2.297188e-03    6.347703e-04    0.001379    1.824648e-03    2.1586
contrast   637.0    3.711644e+01    4.331805e+00    24.254429    3.399963e+01    3.6928
aspect     637.0    1.000000e+00    1.111095e-16    1.000000    1.000000e+00    1.0000

Distribución de clases:
label
malignant    377
benign       260
Name: count, dtype: int64
ESTRUCTURA DE DIRECTORIOS:
  Ruta benigno: C:/Users/USE/Downloads/ProyFathActualizado/ProyFath/data/benign
  Ruta maligno: C:/Users/USE/Downloads/ProyFathActualizado/ProyFath/data/malignar
C:/Users/USE/Downloads/ProyFathActualizado/ProyFath/venv\Lib\site-packages\streaml
fig.savefig(image, **kwargs)
```

DEMO

Login

EDA

Entrenamiento

Home

Prediccion por imagen

Predicción por Imagen

Sube una imagen tiroidea y ejecuta el diagnóstico asistido por IA.

✓ Modelo CNN cargado correctamente

✓ Modelo Random Forest cargado correctamente

Sube una imagen para diagnóstico

Drag and drop file here

Límit 200MB per file • PNG, JPG, JPEG

Browse files

No se subió ninguna imagen; por favor vuelve a ejecutar la celda y carga una imagen.

Drag and drop file here

Browse files

70x460 40,383

The new, os_type, error parameter has been deprecated and will be removed in a future release. Please utilize the error parameter instead.

Imagen no válida

Predicción CNN (fine-tuned): malignant (p=0.596)

Predicción RF: malignant (p=0.825)

Probabilidad combinada malignidad: 0.711

Informe PDF guardado en:

Diagnóstico profesional (visualizado en notebook y guardado en PDF)

Imagen analizada

Informe Asistido por IA - Clínica de Análisis Avanzado de Imágenes (IA)

Fecha: 2025-12-09 16:46:16

Resultado combinado: Malignidad 71.1% / Benigno 28.9%

CNN (fine-tuned): malignant (p=0.596)

RandomForest (embeddings): malignant (p=0.825)

Recomendación clínica:

Sospecha de malignidad: Correlacionar con hallazgos clínicos y considerar biopsia (FNAB) para confirmación histológica.

71.1% maligno

Porcentaje (%) - Malignidad (izq) / Benigno (der)

Login

EDA

Entrenamiento

Home

Prediccion por imagen

Imagen cargada

Predicción CNN (fine-tuned): malignant (p=0.596)

Predicción RF: malignant (p=0.825)

Probabilidad combinada malignidad: 0.711

Informe PDF guardado en:
C:/Users/USE/Downloads/ProyFathActualizado/ProyFath/diagnostico_tiroides_informe_A4_landscape.pdf

Diagnóstico profesional (visualizado en notebook y guardado en PDF)

Imagen analizada

Informe Asistido por IA - Clínica de Análisis Avanzado de Imágenes (IA)

Fecha: 2025-12-09 16:46:16

Resultado combinado: Malignidad 71.1% / Benigno 28.9%

CNN (fine-tuned): malignant (p=0.596)

RandomForest (embeddings): malignant (p=0.825)

Recomendación clínica:

Sospecha de malignidad: Correlacionar con hallazgos clínicos y considerar biopsia (FNAB) para confirmación histológica.

71.1% maligno

Porcentaje (%) - Malignidad (izq) / Benigno (der)

Notas finales:

Se presenta la matriz de confusión del CNN (balanceada 1:1) con colores por cuadrante.

La matriz de correlación muestra pesos (valores numéricos) en cada celda.

Se aplicó fine-tuning simple al CNN y GridSearch al RF; cuadro antes/después mostrado.

El informe profesional se genera en A4 horizontal (guardado como PDF en Drive).

> 📄 Logs del módulo

⬇️ Descargar logs

Evaluación y Resultados

Modelo	Accuracy
CNN (Baseline)	0.6378
CNN (Fine-tuning)	0.6378
Random Forest (Base)	0.6094
Random Forest (Optim.)	0.6406

Resultados comparativos de modelos con entrenamiento base y optimizados

CNN (Baseline & Fine-tuning)

- Accuracy estable: 0.6378
- Fine-tuning no mejora el desempeño
- Sensible al tamaño limitado del dataset

Random Forest (Base)

- Accuracy: 0.6094
- Funciona como baseline tradicional

Random Forest (Optimizado)

- Mejor resultado global: 0.6406
- Mejora gracias a ajuste de hiperparámetros

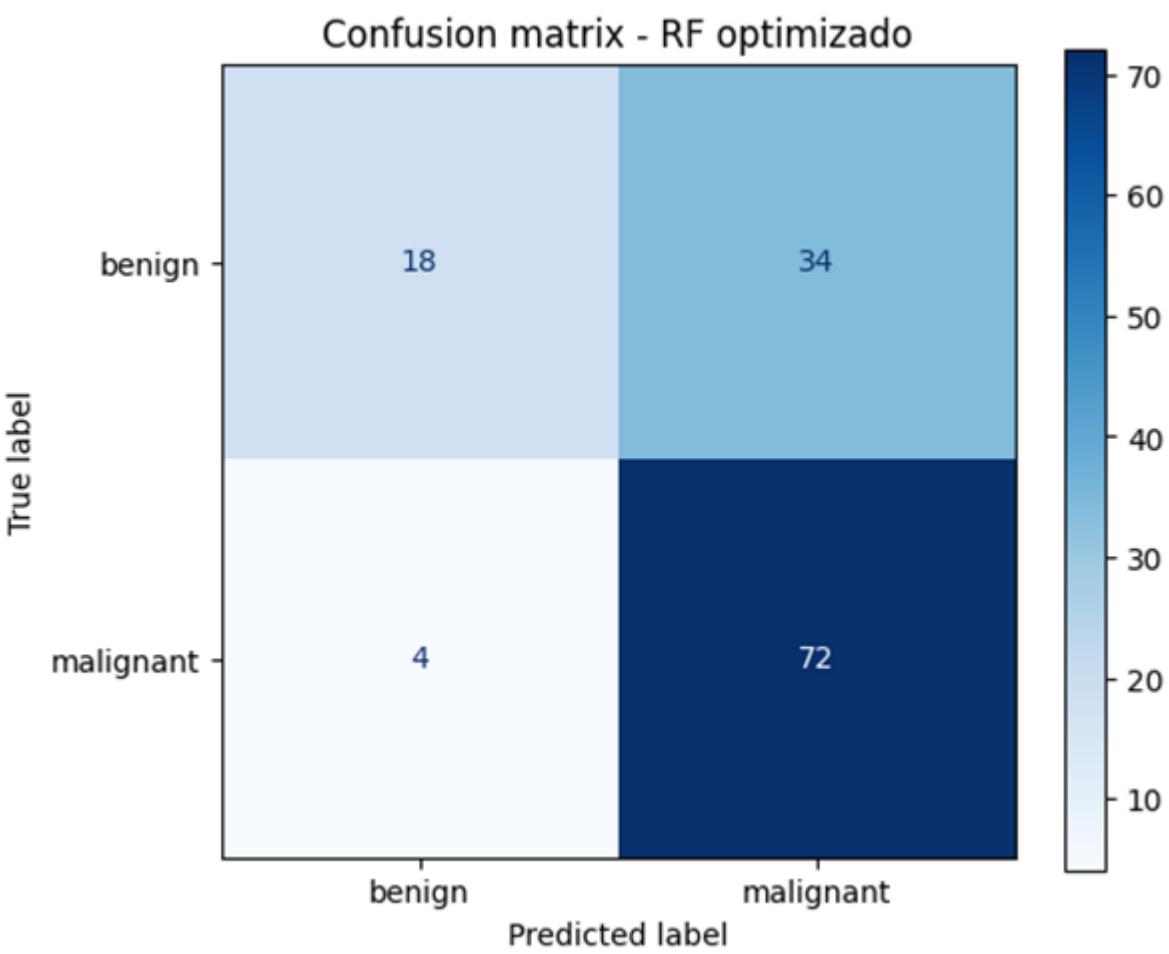
Conclusiones

- Todos los modelos tienen rendimiento similar
- Optimización aporta mejoras ligeras
- CNN podría mejorar con más datos y regularización

Evaluación y Resultados

Comparación, Errores y Casos Clave RF Optimizado

Clase	Precision	Recall	F1-score	Support
benign	0.82	0.55	0.59	52
malignant	0.68	0.95	0.79	76
Accuracy			0.70	128
Macro avg	0.75	0.65	0.64	128
Weighted avg	0.74	0.70	0.67	128



Consideraciones Éticas

1. Sesgos identificados

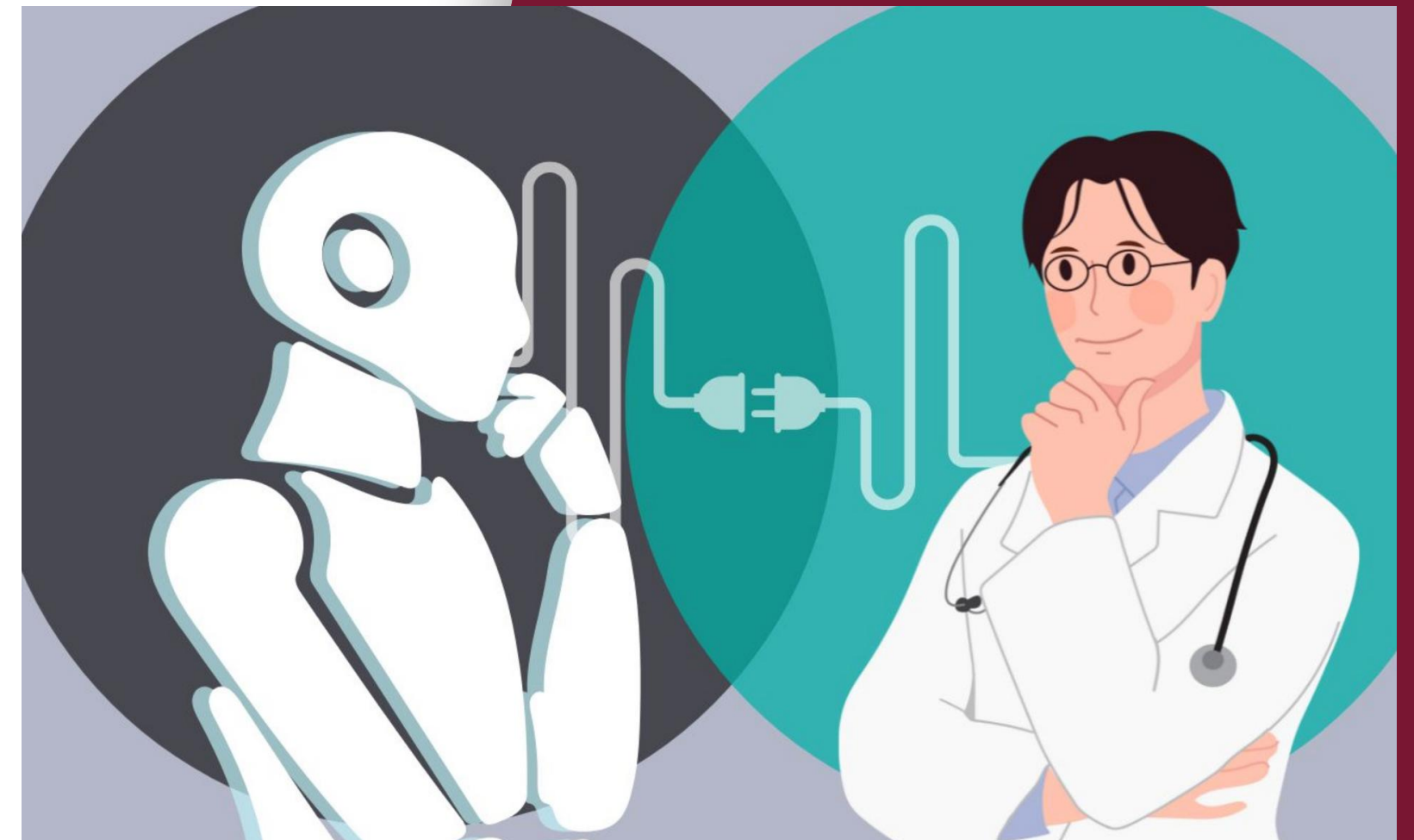
Posible sesgo por procedencia y calidad desigual de las imágenes del dataset.

Representación limitada de ciertos grupos clínicos o etarios.

2. Impacto social potencial

Contribuye a la detección temprana del cáncer de tiroides y optimiza la atención médica.

Riesgo de malinterpretación si se usa sin supervisión profesional.



Consideraciones Éticas

3. Medidas de mitigación

Uso de datos anónimos y públicos.

Validación cruzada para reducir sesgos y errores.

Definido como herramienta de apoyo, no sustituto del médico.

4. Limitaciones éticas

Dependencia de la calidad y diversidad de los datos.

Requiere validación clínica adicional antes de aplicarse en hospitales.

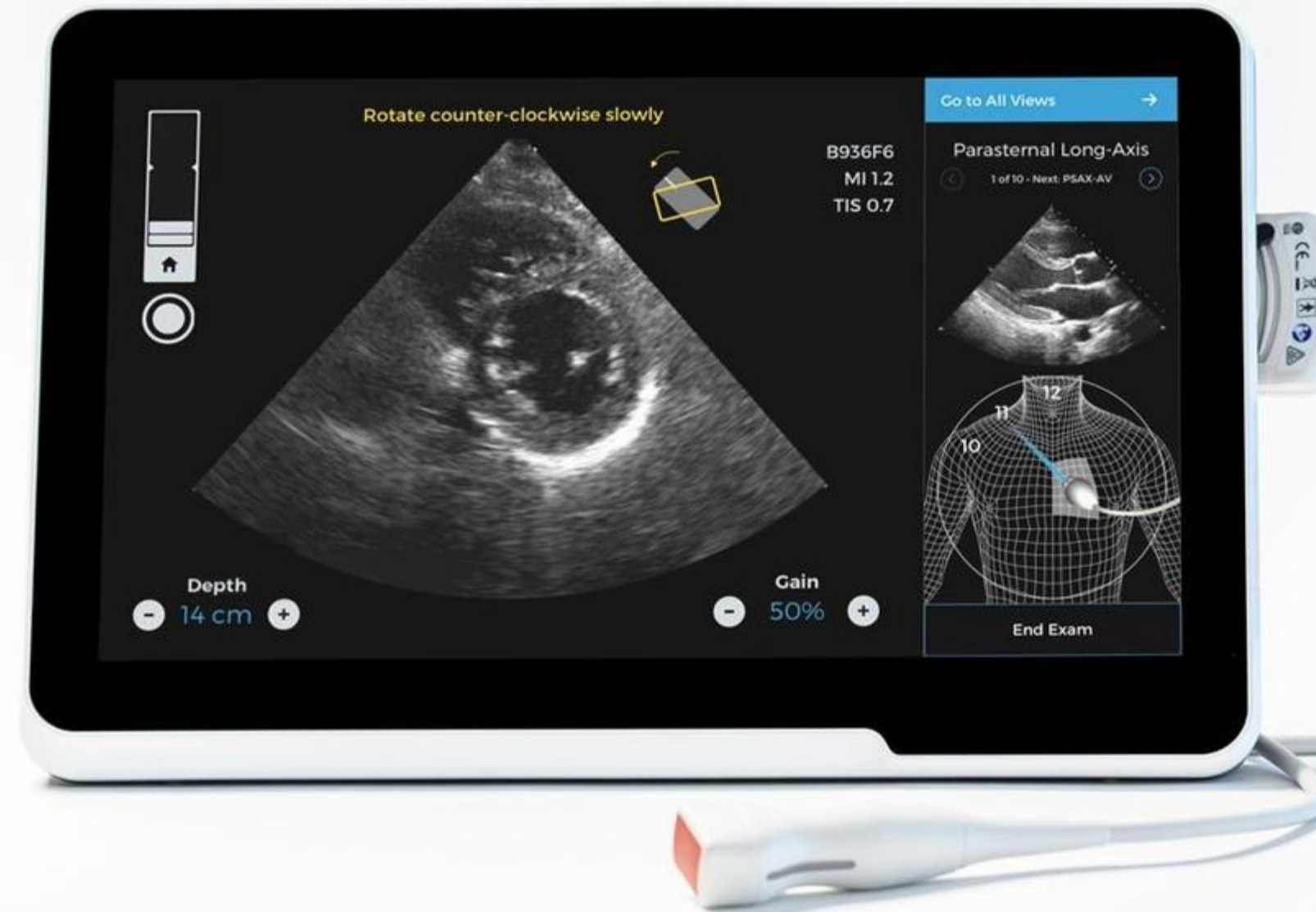
Necesario mantener intervención humana en la toma de decisiones.



Conclusiones y Trabajo Futuro

Logros principales

- Sistema capaz de clasificar nódulos tiroideos benignos y maligno.
- Preprocesamiento, modelado, evaluación y almacenamiento del modelo entrenado.
- Viabilidad del uso de IA.



Lecciones aprendidas

- La calidad del dataset es determinante.
- La interpretabilidad y validación clínica son tan importantes.
- La integración entre IA y medicina requiere colaboración interdisciplinaria constante.

Conclusiones y Trabajo Futuro

Posibles mejoras futuras

- Ampliar el dataset.
- Implementar mas imágenes para que el modelo CNN sea mejor.
- Incluir un módulo de explicabilidad.

Aplicaciones adicionales

- Extender el enfoque.
- Desarrollar una interfaz web médica implementada a HIS.
- Integrar en sistemas hospitalarios.

