

# **Reto DaCiberSalut**

## **Generación de modelos IA para identificación de tejidos**

Iván Falcón Monzón, Ismael Díaz Sancha, David Rodríguez Gerrard

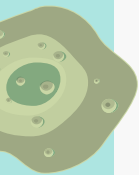
CEIABD | 2025

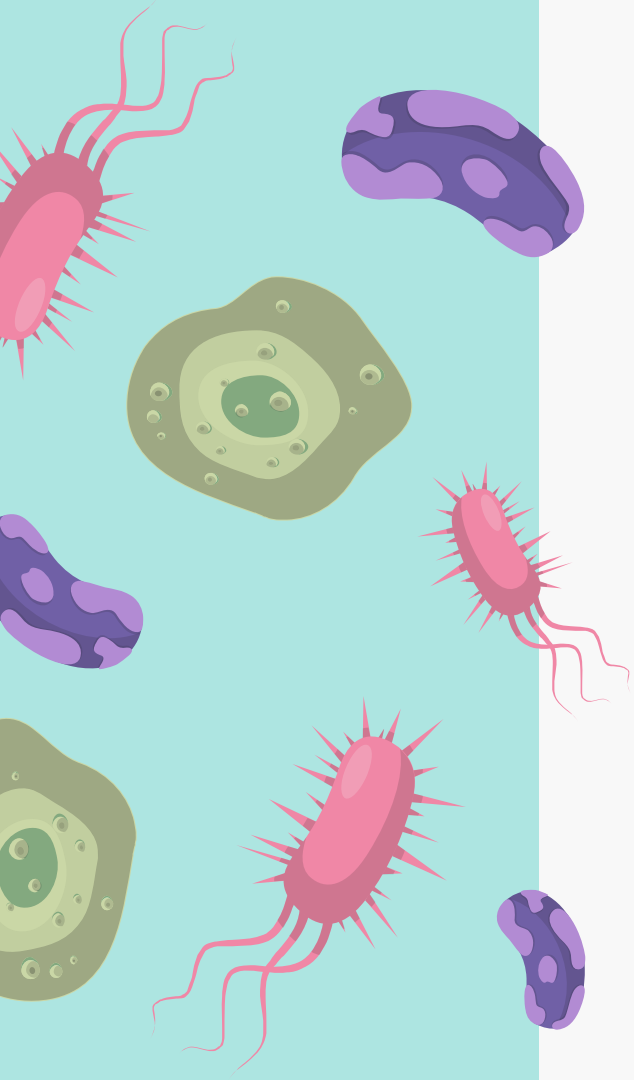




# ÍNDICE

1. **Objetivo del Proyecto**
2. **Tecnologías utilizadas**
3. **Estructura del proyecto**
4. **Trabajo en equipo y roles por fase**
5. **Dificultades encontradas y cómo las resolvimos**
6. **Conclusión y posibles mejoras**





**01**

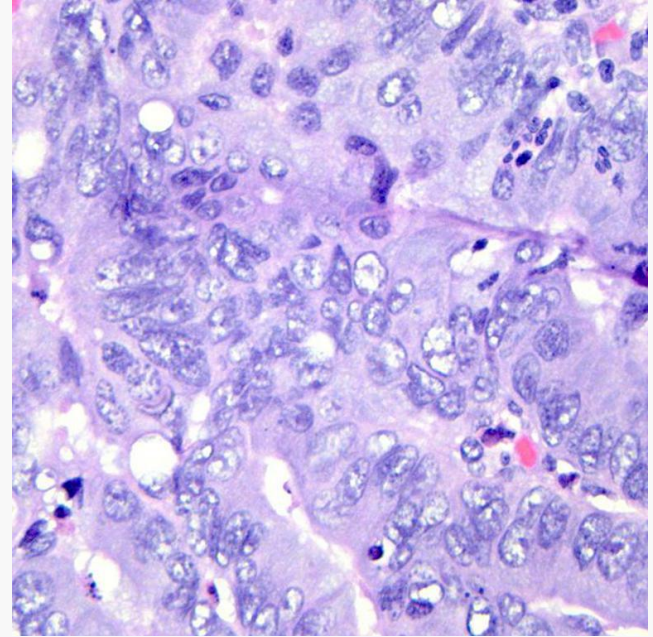
# Objetivo del Proyecto

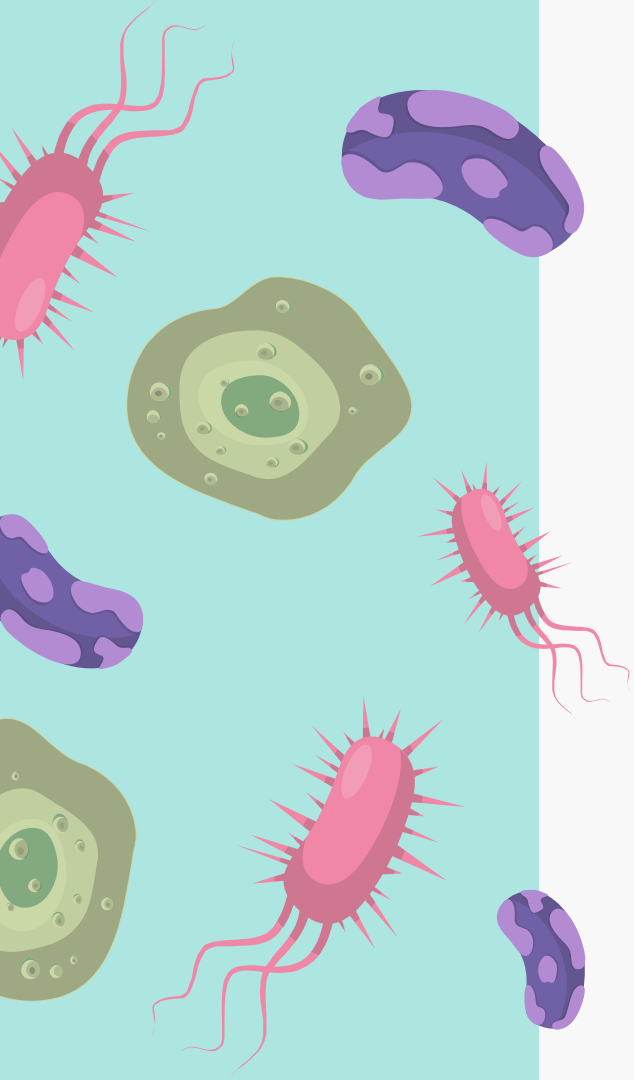


# Objetivo

## ¿Qué problema queremos resolver?

“Diseñar una solución tecnológica que permita gestionar y visualizar datos médicos de forma intuitiva para profesionales.”





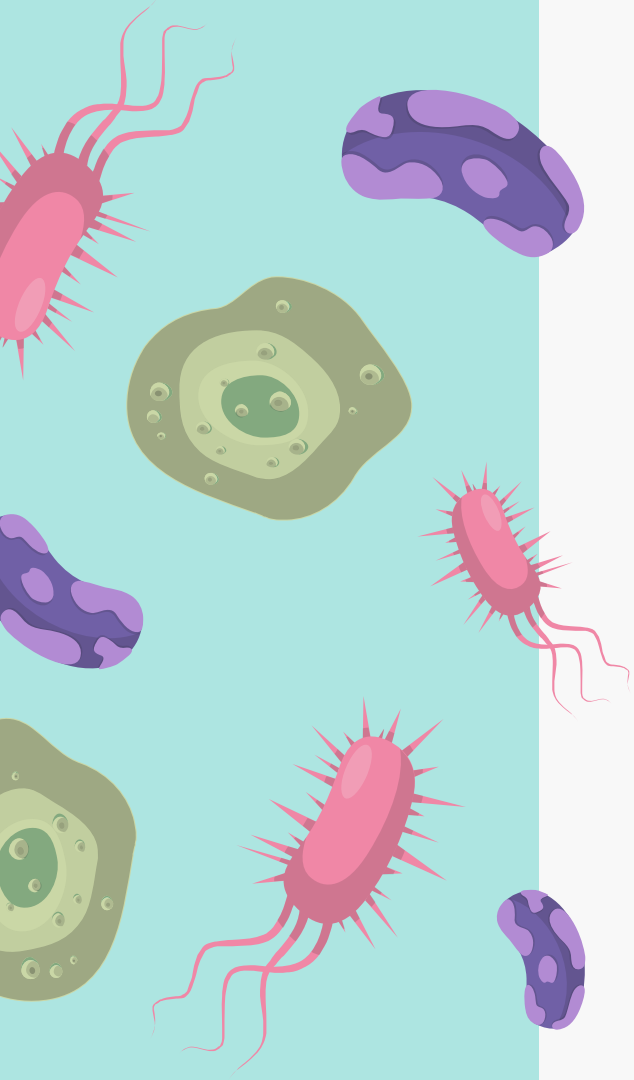
02

## Tecnologías utilizadas



# Tecnologías utilizadas





**03**

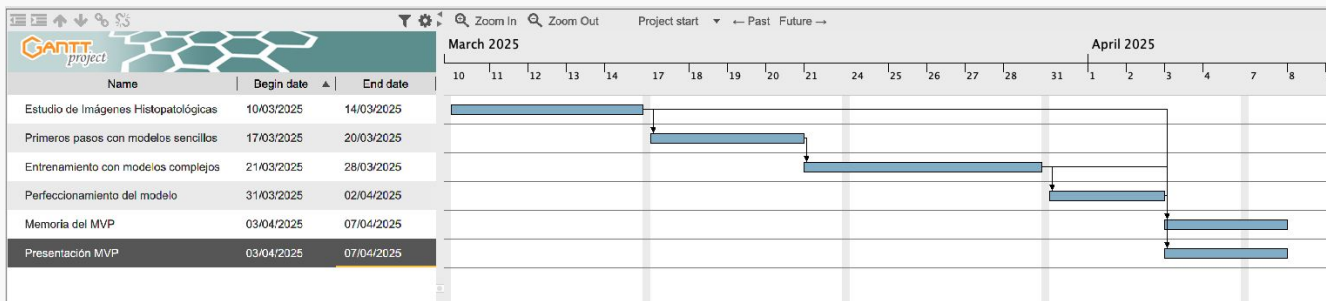
## **Fases del proyecto**



# Fases del proyecto

## Estructura

El **objetivo del proyecto** fue evaluar la viabilidad de entrenar modelos de aprendizaje automático, utilizando recursos accesibles para estudiantes, con el fin de detectar indicios de cáncer en imágenes histopatológicas. A pesar de emplear un conjunto de datos reducido y limitado en clase, los resultados evidencian la alta capacidad generalizadora de los modelos y su potencial aplicabilidad en entornos educativos y clínicos.





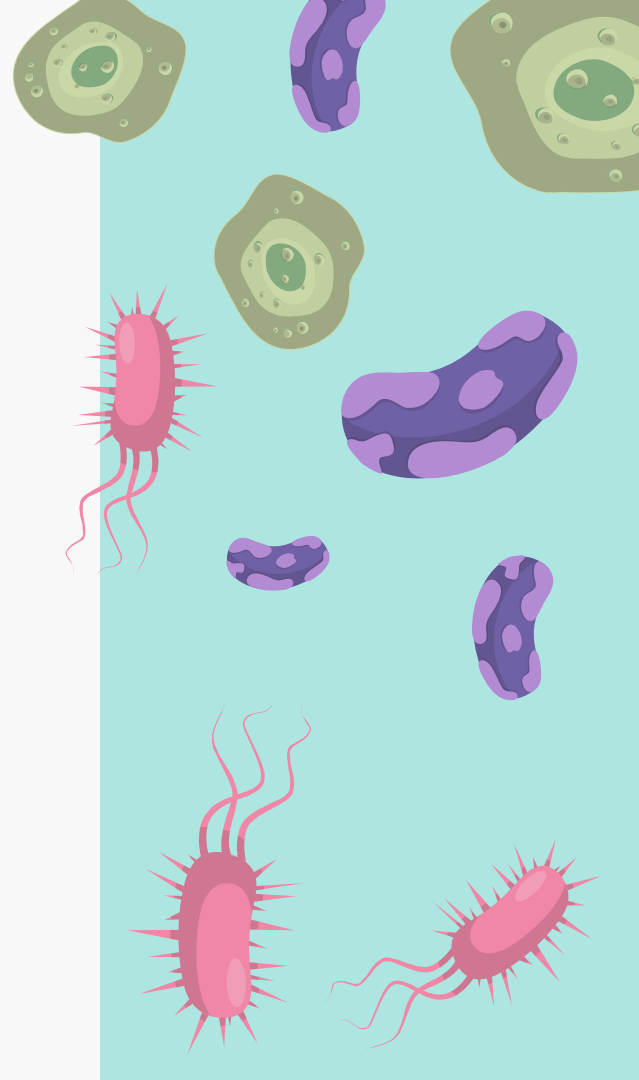
# Fases Iniciales

## Investigación inicial

El proyecto se desarrolló en cinco fases principales. En la fase de **investigación inicial**, se llevó a cabo una revisión de literatura centrada en la clasificación histológica mediante imágenes y el uso de inteligencia artificial en tareas de diagnóstico médico. También se evaluaron distintas tecnologías y herramientas accesibles para estudiantes, con el fin de garantizar la viabilidad del desarrollo en un entorno académico.

## Procesamiento de datos

Durante la fase de procesamiento de datos, se trabajó con el dataset **LC25000**, compuesto por imágenes histopatológicas del pulmón y el colon, abarcando cinco clases: *colon\_aca*, *colon\_n*, *lung\_aca*, *lung\_n* y *lung\_scc*. Las imágenes fueron reorganizadas desde su estructura original hacia un único directorio estructurado por clases, compatible con el formato ImageFolder de PyTorch. Posteriormente, todas las imágenes fueron redimensionadas uniformemente a **512x512 píxeles** y **normalizadas utilizando los valores estándar de ImageNet**, asegurando coherencia para su uso con modelos preentrenados. Además, se realizó una división aleatoria del conjunto de datos en subconjuntos de entrenamiento (80%) y validación (20%), respetando la distribución de clases original.



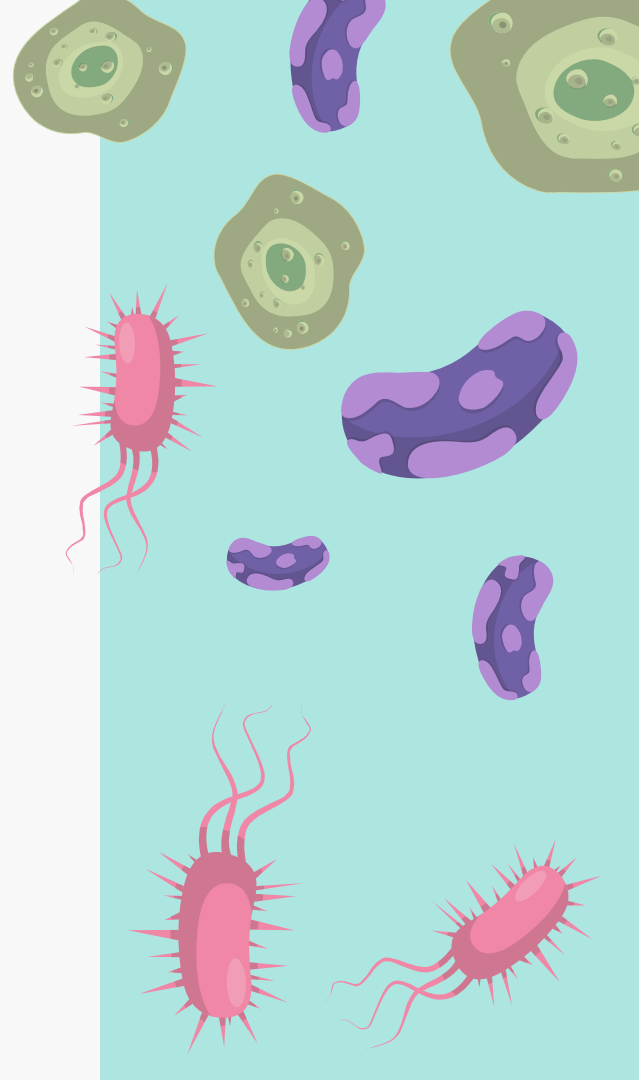
# Fase Intermedia

## Desarrollo del modelo de IA

La fase de desarrollo del modelo se centró en la **implementación de un modelo basado en transferencia de aprendizaje**, utilizando la arquitectura **ResNet50** preentrenada en ImageNet. Se congelaron todas las capas convolucionales del modelo y se reemplazó la capa final para adaptarla a las cinco clases del dataset. El entrenamiento se llevó a cabo en un entorno con soporte para GPU, utilizando el optimizador Adam y la función de pérdida CrossEntropyLoss. Aunque no se aplicaron técnicas explícitas de aumento de datos (*data augmentation*), se consideraron como una futura mejora. Finalmente, se evaluó el modelo en un conjunto de validación independiente, y se guardó el modelo entrenado en formato .pth para su posterior reutilización en tareas de inferencia.

## Implementación de la aplicación web

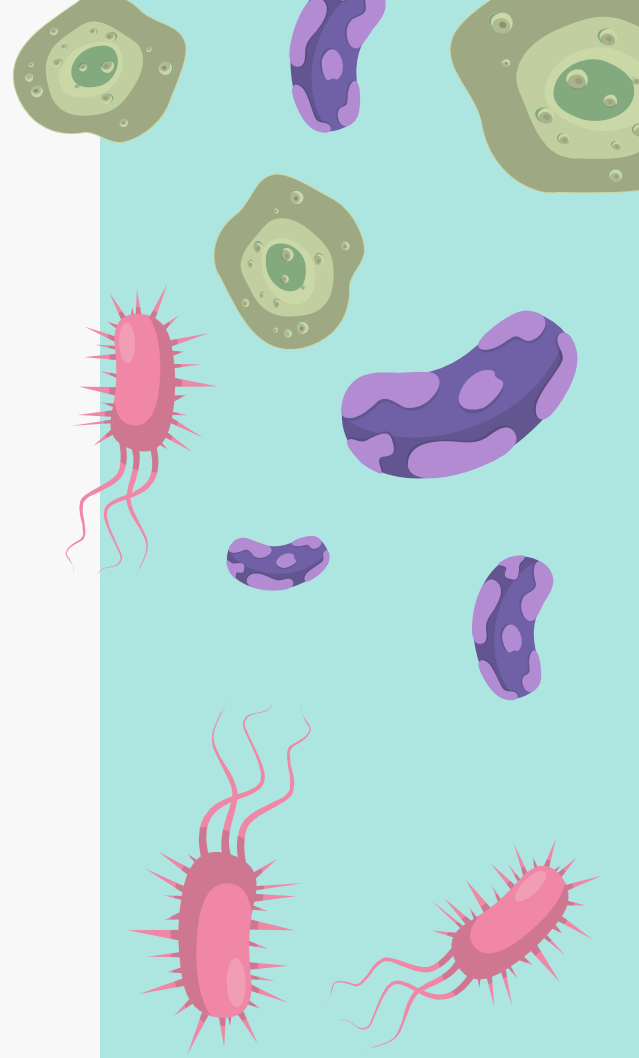
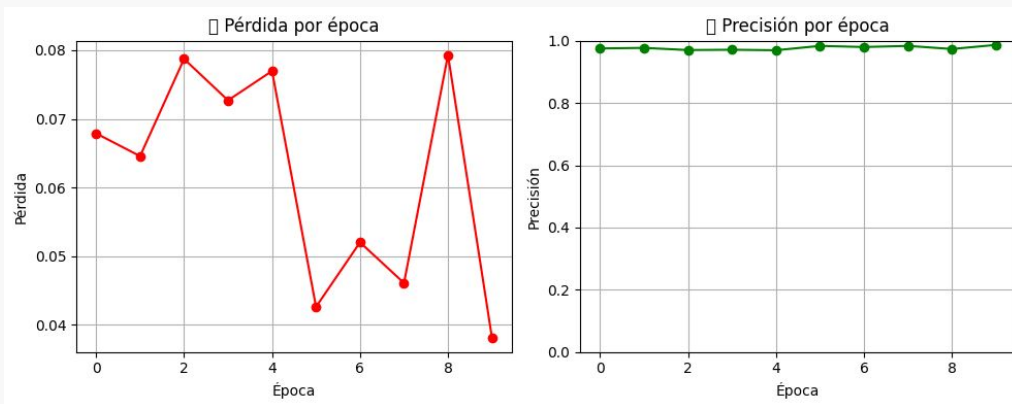
En la **implementación de la aplicación web**, se creó una interfaz sencilla utilizando HTML, CSS, JavaScript y Flask. Esta interfaz permite al usuario cargar imágenes histológicas y obtener los resultados del modelo IA, los cuales se muestran de manera estructurada en una tabla, facilitando su interpretación.

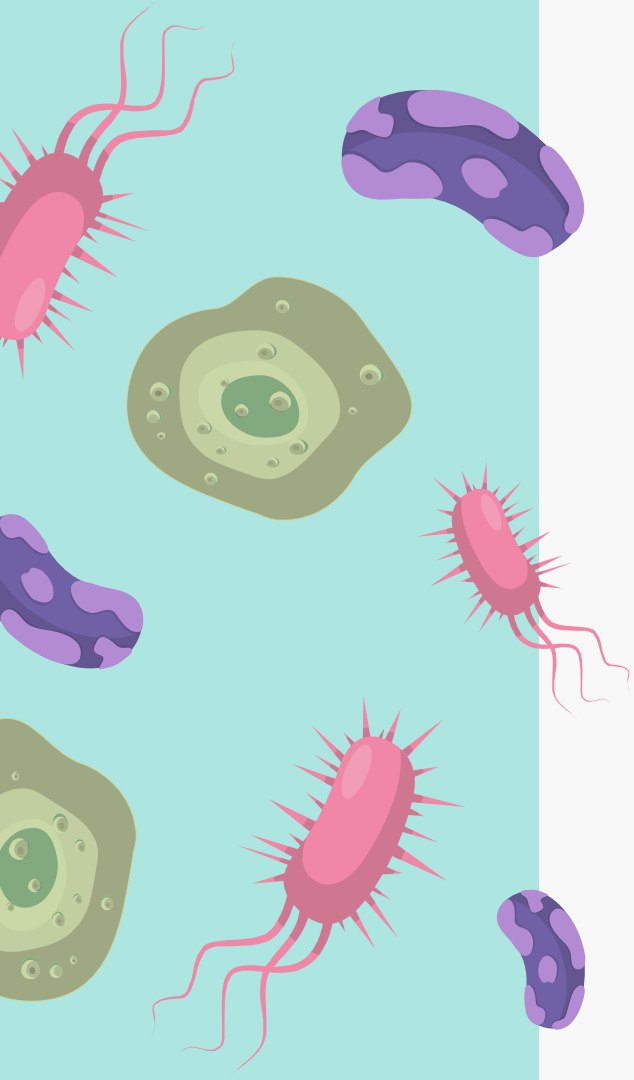


# Fase Final

## Evaluación y mejoras

Finalmente, la fase de **evaluación y mejoras** incluyó pruebas con imágenes que el modelo no había visto previamente, permitiendo analizar tanto los aciertos como los errores en las predicciones. A partir de esta evaluación, se propusieron mejoras en la visualización de resultados para hacer la herramienta más comprensible y útil.





# 04

## Trabajo en equipo y roles



# Trabajo en equipo y roles por fases

## **Fase 1: Investigación** (2 semanas)

Definición del problema, objetivos y análisis de casos similares.

**David** lideró esta fase con un fuerte enfoque en la investigación inicial.

## **Fase 2: Preparación de datos** (2 semanas)

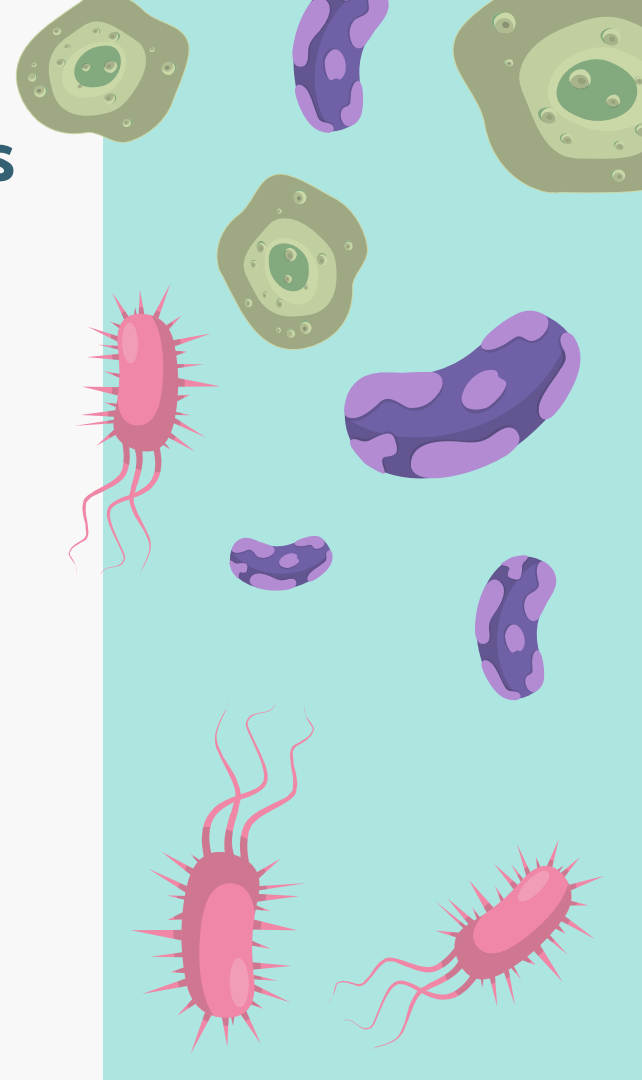
Obtención, limpieza y normalización del dataset (LC25000).

Trabajo conjunto entre los tres miembros.

## **Fase 3: Desarrollo del modelo** (3 semanas)

Diseño y entrenamiento de una red neuronal convolucional.

**Ismael** asumió el liderazgo técnico en esta fase.



# Trabajo en equipo y roles por fases

## **Fase 4: Desarrollo de la app web** (3 semanas)

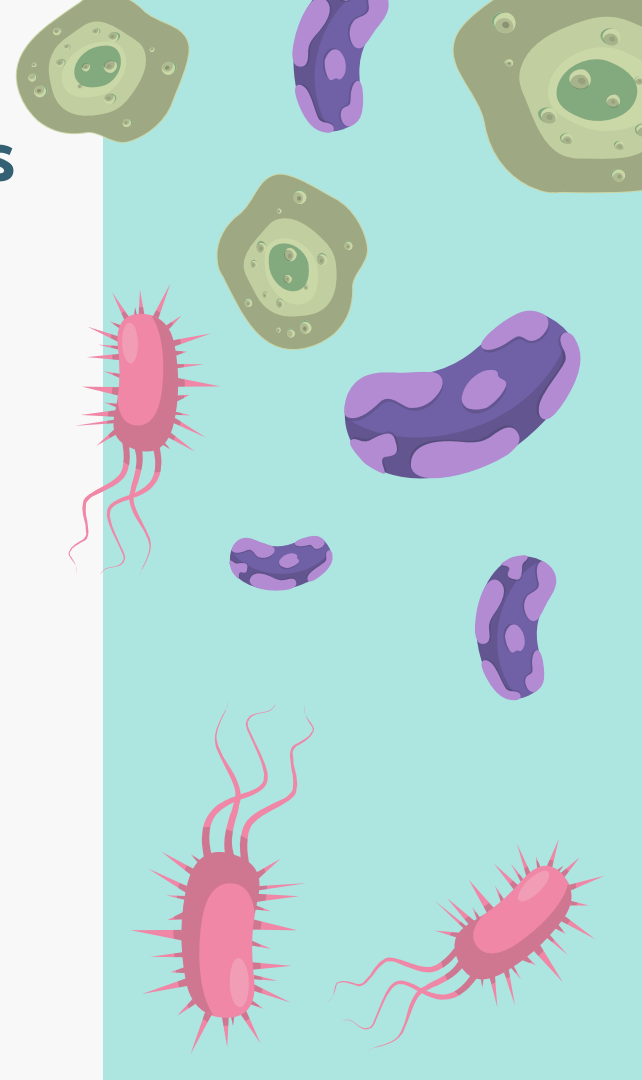
Creación de una interfaz visual e integración con el modelo IA.

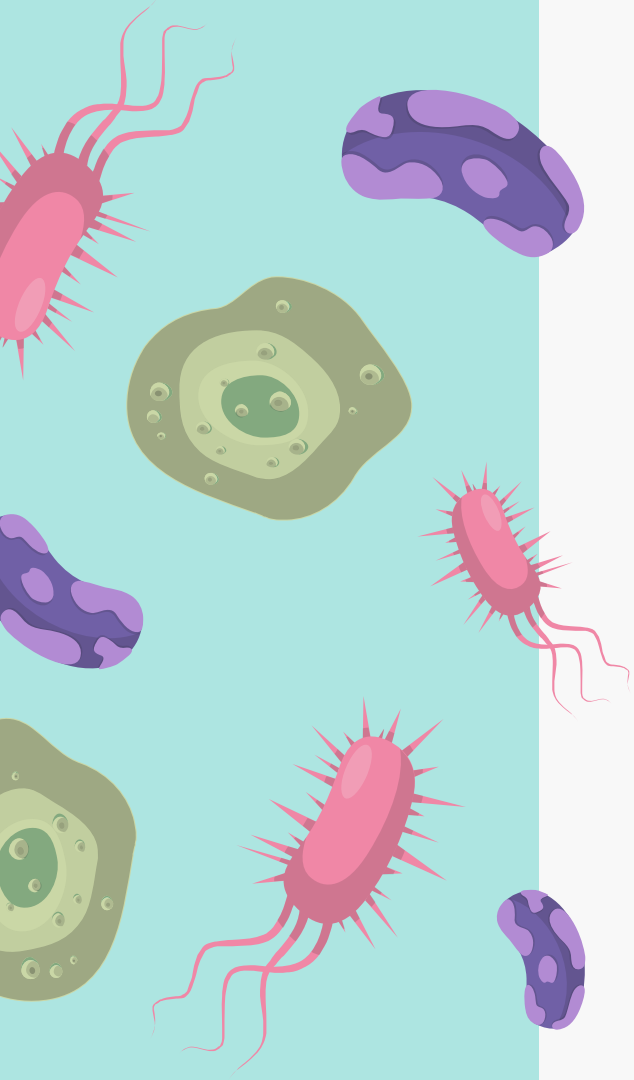
**Iván** lideró esta etapa, enfocándose en la experiencia del usuario.

## **Fase 5: Evaluación y documentación** (2 semanas)

Pruebas finales, análisis de resultados y redacción de la memoria.

Esta fase fue trabajada de forma conjunta por todo el equipo.





# 05

**Dificultades  
encontradas y cómo  
las resolvimos**



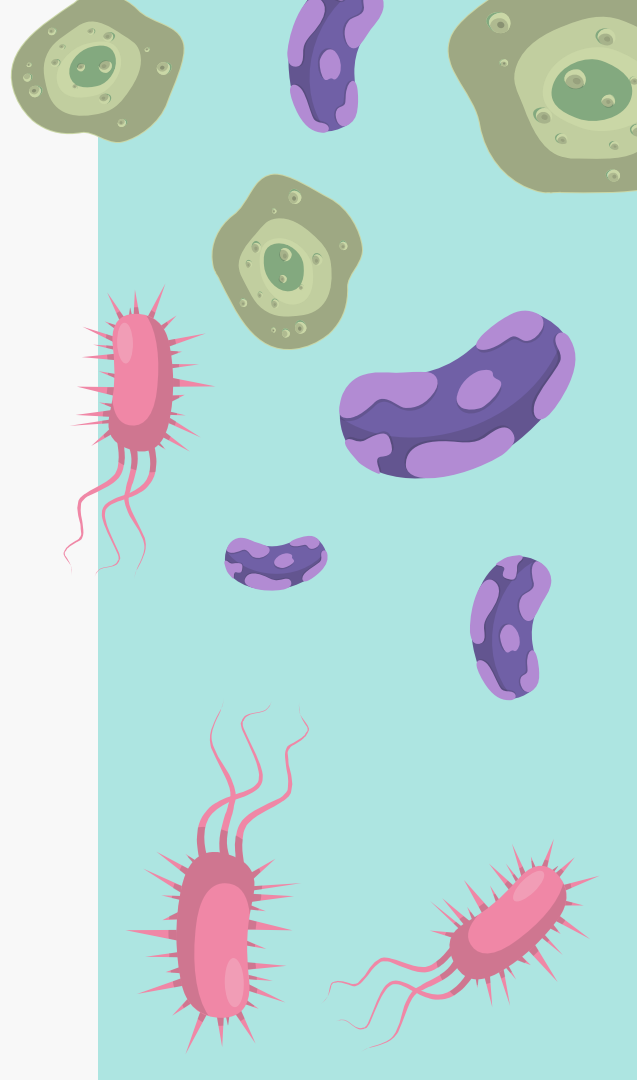
# Dificultades encontradas

## Dificultades

**Elección del modelo:** Decidir qué arquitectura usar fue un reto. Nos decidimos por **ResNet**, aunque había opciones como EfficientNet que también eran viables.

**Clasificación de imágenes no válidas:** El modelo tendía a intentar clasificar imágenes que no eran histológicas.

**Visualización de resultados:** Mostrar los resultados de forma comprensible fue un desafío, especialmente para usuarios sin conocimientos técnicos.





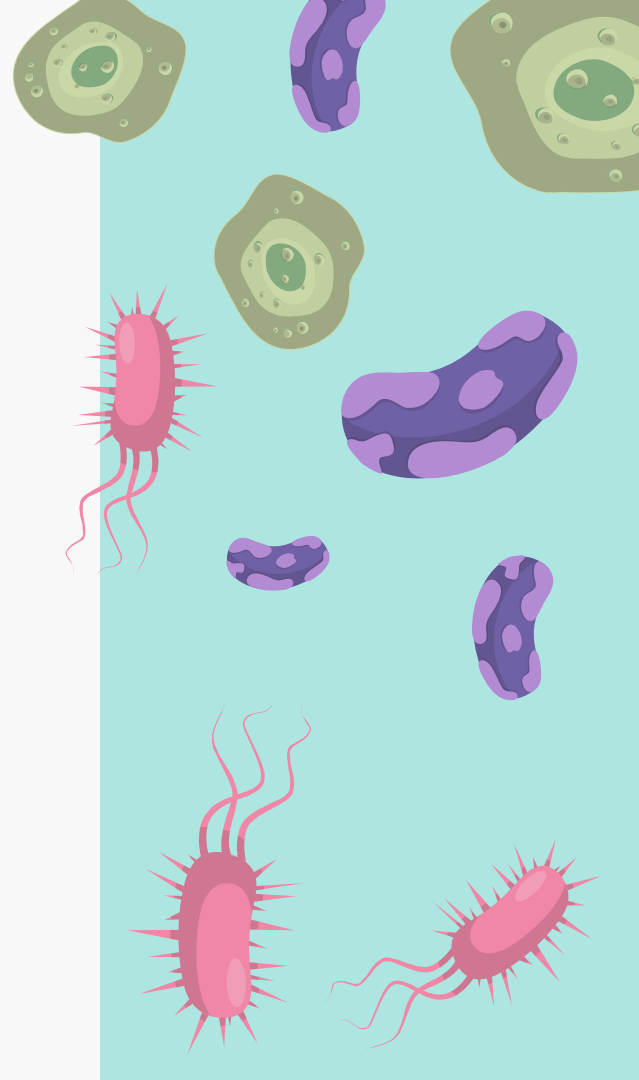
# Como las resolvimos

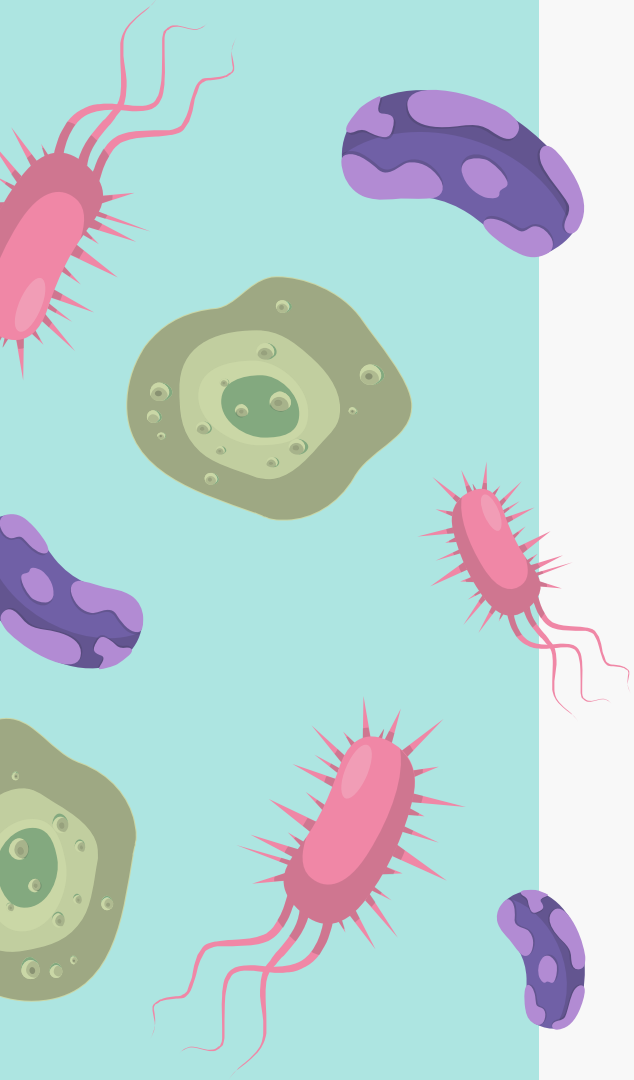
## Soluciones

**Modelo ResNet:** Se eligió por su estabilidad y buen desempeño en tareas de clasificación profunda.

**Filtrado y control de imágenes:** Se trabajó en mejorar la validación de entrada para evitar clasificaciones erróneas.

**App visual con flask:** Se desarrolló una interfaz simple y clara, facilitando la interacción y comprensión de los resultados por parte del usuario.





**06**

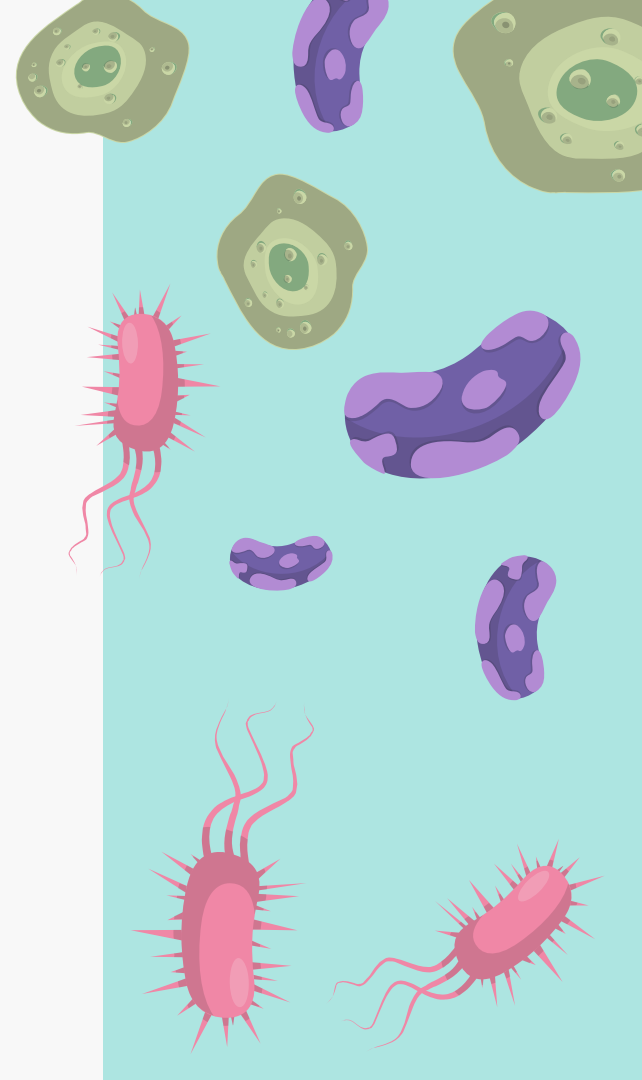
## **Conclusión y posibles mejoras**



# Conclusiones y Posibles Mejoras

## Conclusiones

El objetivo del proyecto fue evaluar la viabilidad de entrenar modelos de aprendizaje automático, utilizando recursos accesibles para estudiantes, con el fin de detectar indicios de cáncer en imágenes histopatológicas. A pesar de emplear un conjunto de datos reducido y limitado en clases, los resultados evidencian la alta capacidad generalizadora de los modelos y su potencial aplicabilidad en entornos educativos y clínicos.



# Conclusiones y Posibles Mejoras

## Posibles mejoras

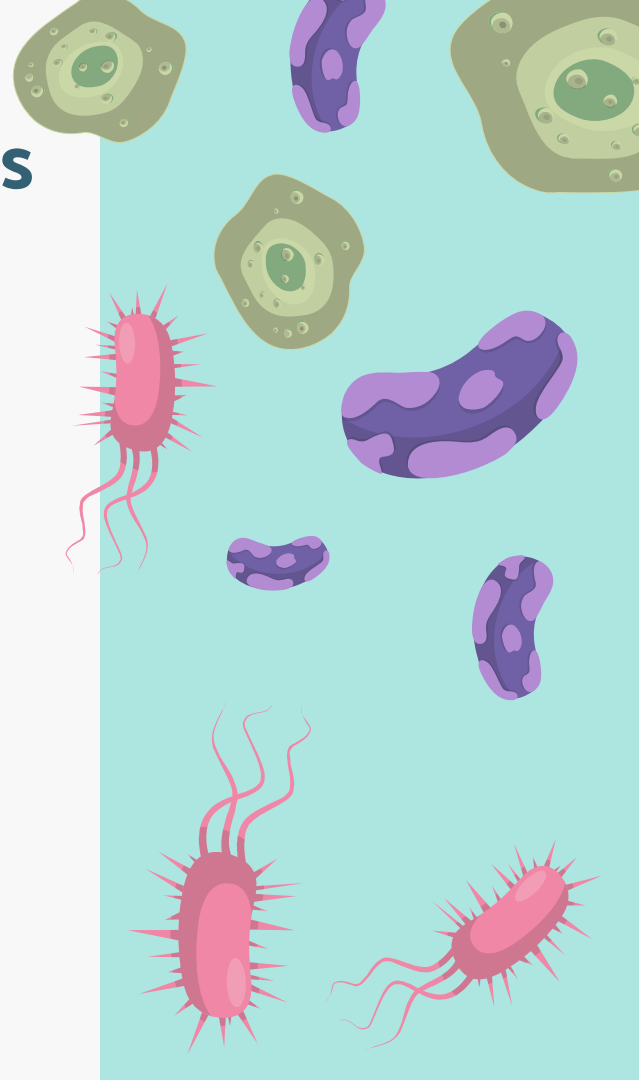
**Ampliar y diversificar los datos:** Incorporar más imágenes y clases para mejorar la capacidad de generalización del modelo.

**Optimizar el modelo:** Probar diferentes arquitecturas, usar modelos pre entrenados y ajustar hiper parámetros clave.

**Aumentar los datos (Data Augmentation):** Aplicar transformaciones a las imágenes para enriquecer el dataset sin necesidad de nuevos datos reales.

**Mejorar la evaluación:** Usar validación cruzada y métricas más completas como sensibilidad, especificidad y AUC.

**Automatizar y documentar el proceso:** Crear un pipeline reproducible para facilitar mejoras futuras e integración en aplicaciones reales.





# Webgrafía

## **PyTorch documentation**

<https://pytorch.org/docs/stable/index.html>

## **Resnet Architecture Explained:**

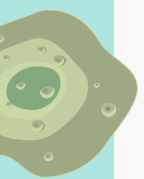
<https://medium.com/@siddheshb008/resnet-architecture-explained-47309ea9283d>

## **Repositorio completo Github:**

[https://github.com/IvanFalconMonzon/DaCiberSalut\\_Iv-n-Ismael-David.git](https://github.com/IvanFalconMonzon/DaCiberSalut_Iv-n-Ismael-David.git)

## **Google Colab (modelo):**

[https://colab.research.google.com/drive/1dpIU8toZaO76nQ31E216yd-wLktG8\\_\\_P?usp=sharing](https://colab.research.google.com/drive/1dpIU8toZaO76nQ31E216yd-wLktG8__P?usp=sharing)



# FIN

Muchas gracias por la atención!

## Clasificador de Imágenes - DaCiberSalut

Nota: Este modelo funciona con los siguientes tipos de tejidos:

Colon Adenocarcinoma (colon\_aca)  
Colon Benign Tissue (colon\_n)  
Lung Adenocarcinoma (lung\_aca)  
Lung Benign Tissue (lung\_n)  
Lung Squamous Cell Carcinoma (lung\_scc)



Seleccionar Imagen

