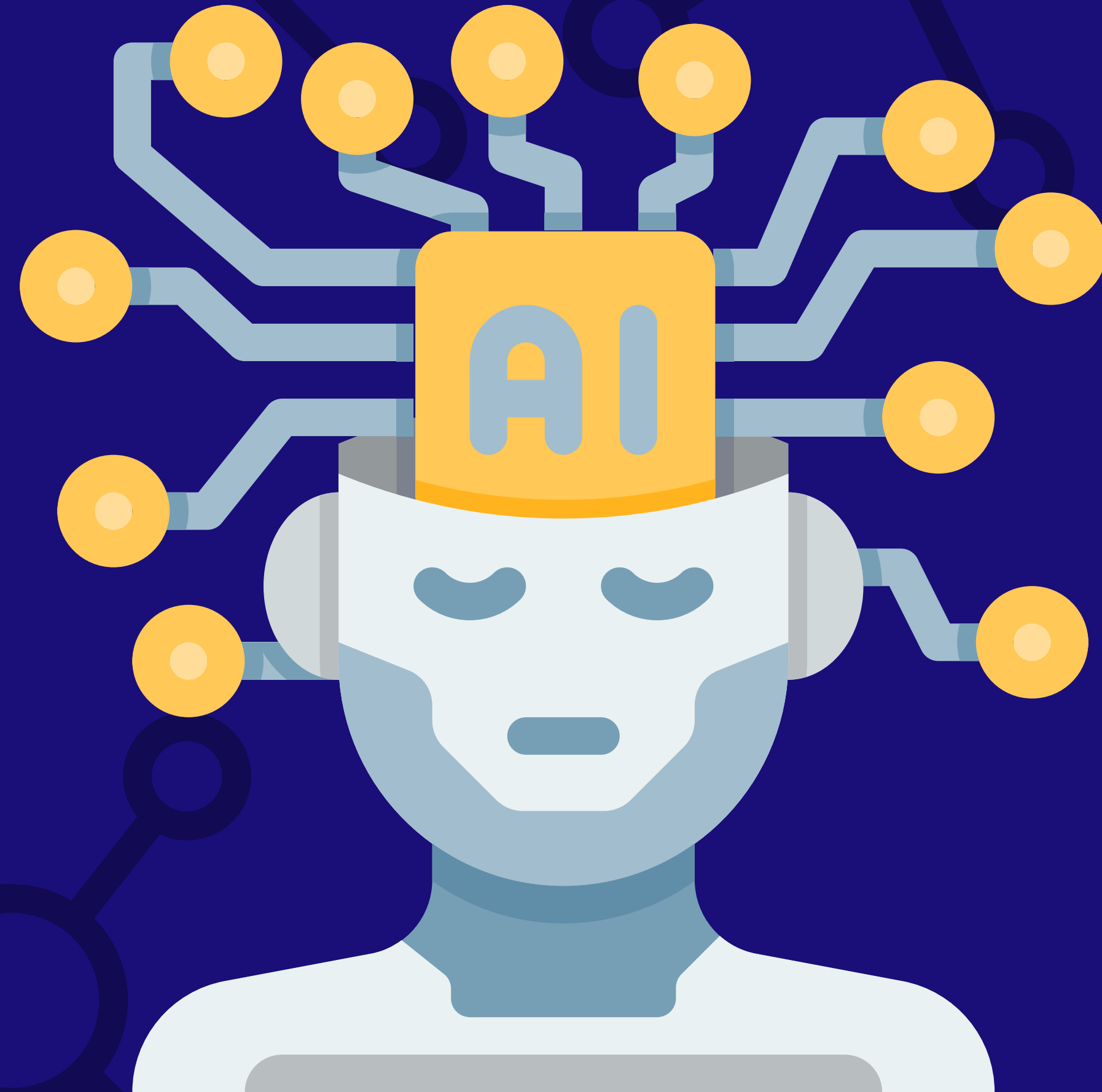


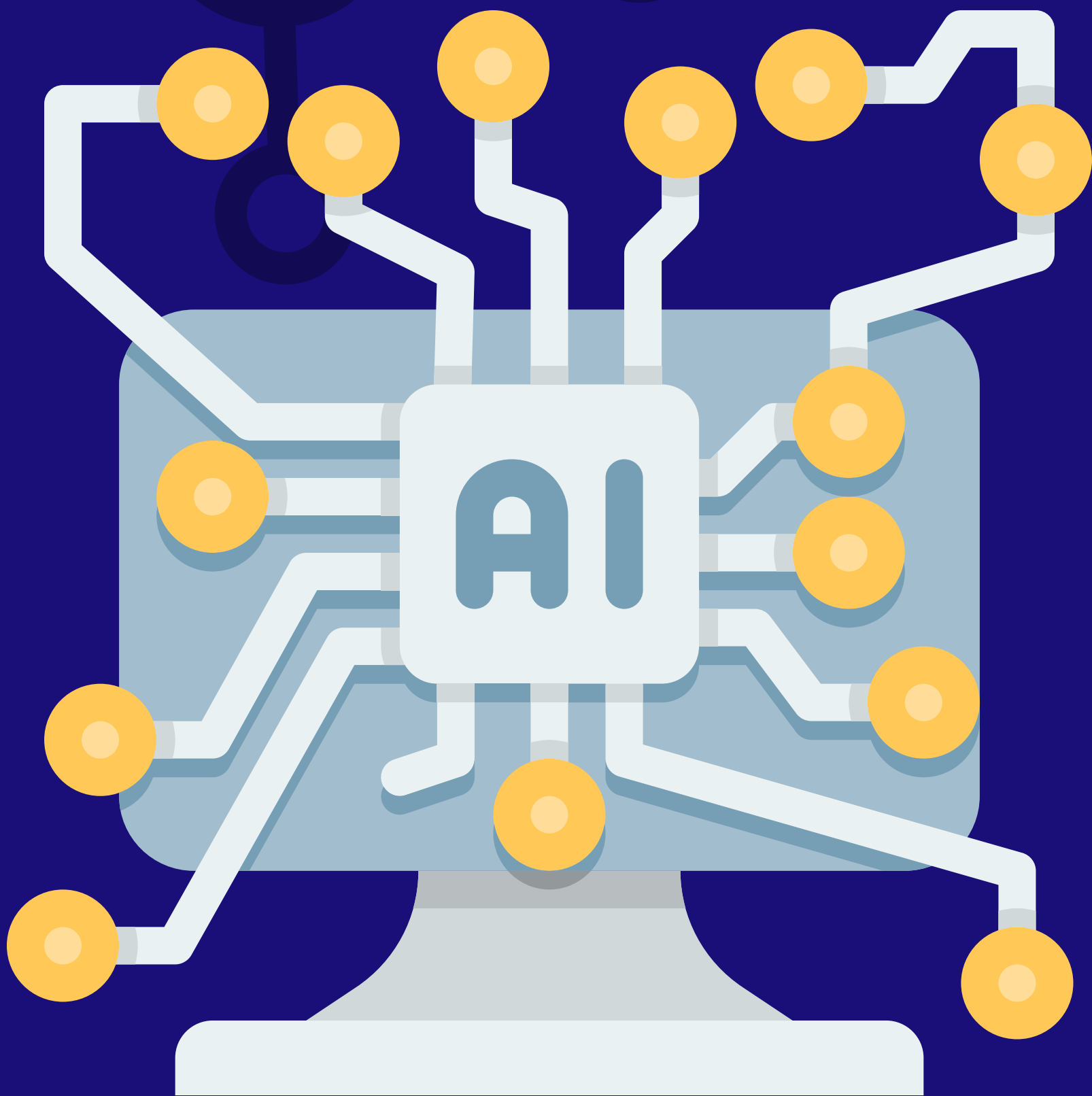


Segmentación de Infecciones Pulmonares con U-Net + EfficientNetB2

TC3006C - Eq. 5

A01027715 - Ulises Orlando Carrizalez Lerin
A01028008 - Tomás Pérez Vera
A01799609 - Bárbara Paola Alcántara Vega
A01710965 - Mónica Monserrat Martínez Vásquez
A01705840 - María José Soto Castro



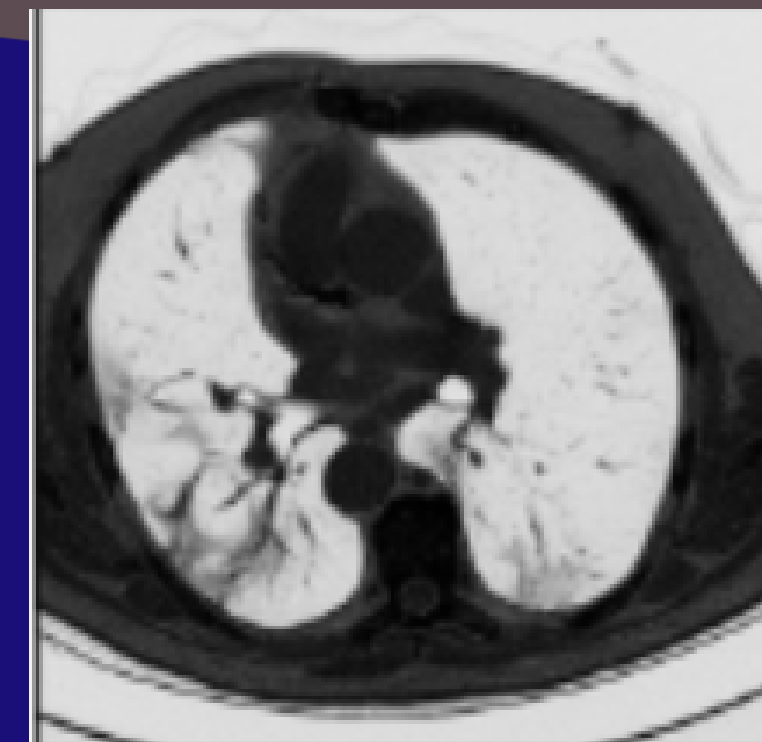


Índice

- 01.** Introducción
- 02.** Análisis de datos
- 03.** Modelo
- 04.** Entrenamiento
- 05.** Resultados
- 06.** Conclusiones

Introducción

Usando imágenes de áreas del pulmón de escaneos de tomografía computada (TC), se quiere segmentar en categorías (pulmón, fluidos, engrosamiento de tejido o fondo de la imagen) para facilitar el diagnóstico del virus de COVID-19.



1

Las imágenes permiten a los doctores saber la severidad de la enfermedad.

2

Se utiliza U-Net y FPN comunmente en imágenes médicas

3

El análisis de imágenes debe de usarse como una herramienta no unicamente para diagnósticos

Objetivo Reto

Proveer evidencia visual de las áreas pulmonares afectadas, mejorando la toma de decisiones clínicas y el tratamiento médico.

Análisis de datos:

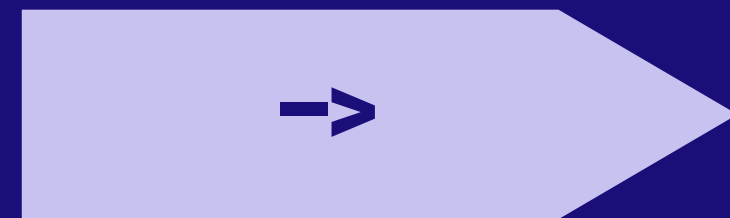


Los datos son originalmente de **SIRM (Società Italiana di Radiologia Medica e Interventistica)**, quienes cuentan con regulaciones éticas ya aplicadas para proteger la confidencialidad de los pacientes y permitiendo su uso académico.

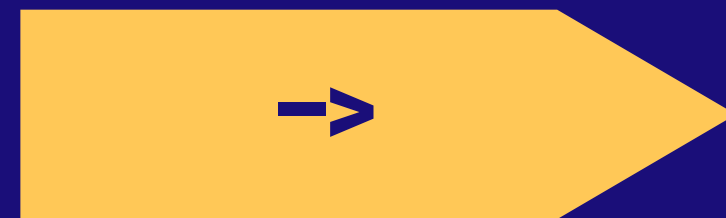
Exploración de datos

Medseg: 100 imágenes con sus máscaras correspondientes en 4 canales:

- Ground glass opacities
- Consolidations
- Unaffected lung tissue
- Background



Radiopedia: 829 capas de CT con la misma estructura que Medseg e imágenes de 512x512 píxeles



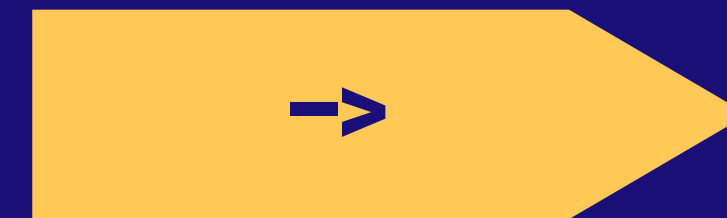
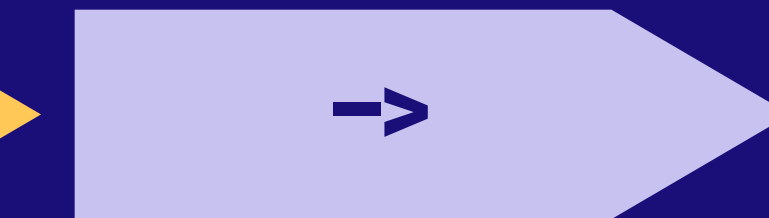
Carga de datos

Los datos (imágenes) son extraídas de sus repositorios y almacenados en arreglos numpy. Las imágenes ya cuentan con formato y sus máscaras como etiquetas.

Preprocesamiento

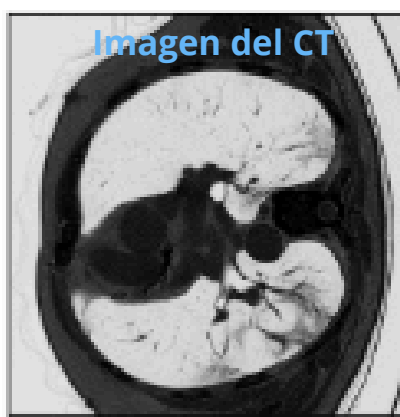
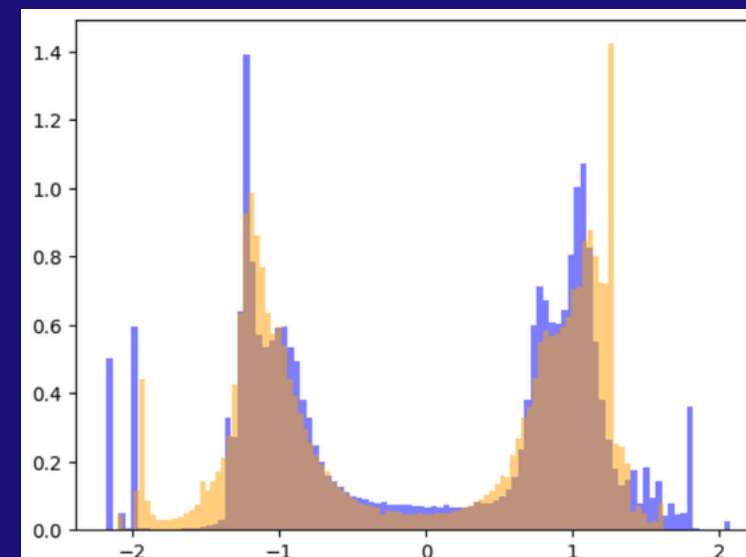
Antes de entrenar:

- Clipping
- Escalamiento robusto
- Estandarización (Z-Score)
- Aumentación

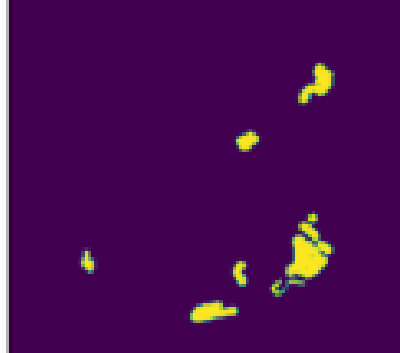


División Train/Val

El set de entrenamiento ya es proporcionado por la convocatoria. Se seleccionaron 24 imágenes de Medseg para el set de validación.



Ground glass opacities



Consolidations



Unaffected lung tissue



Background



Arquitectura utilizada

U-NET

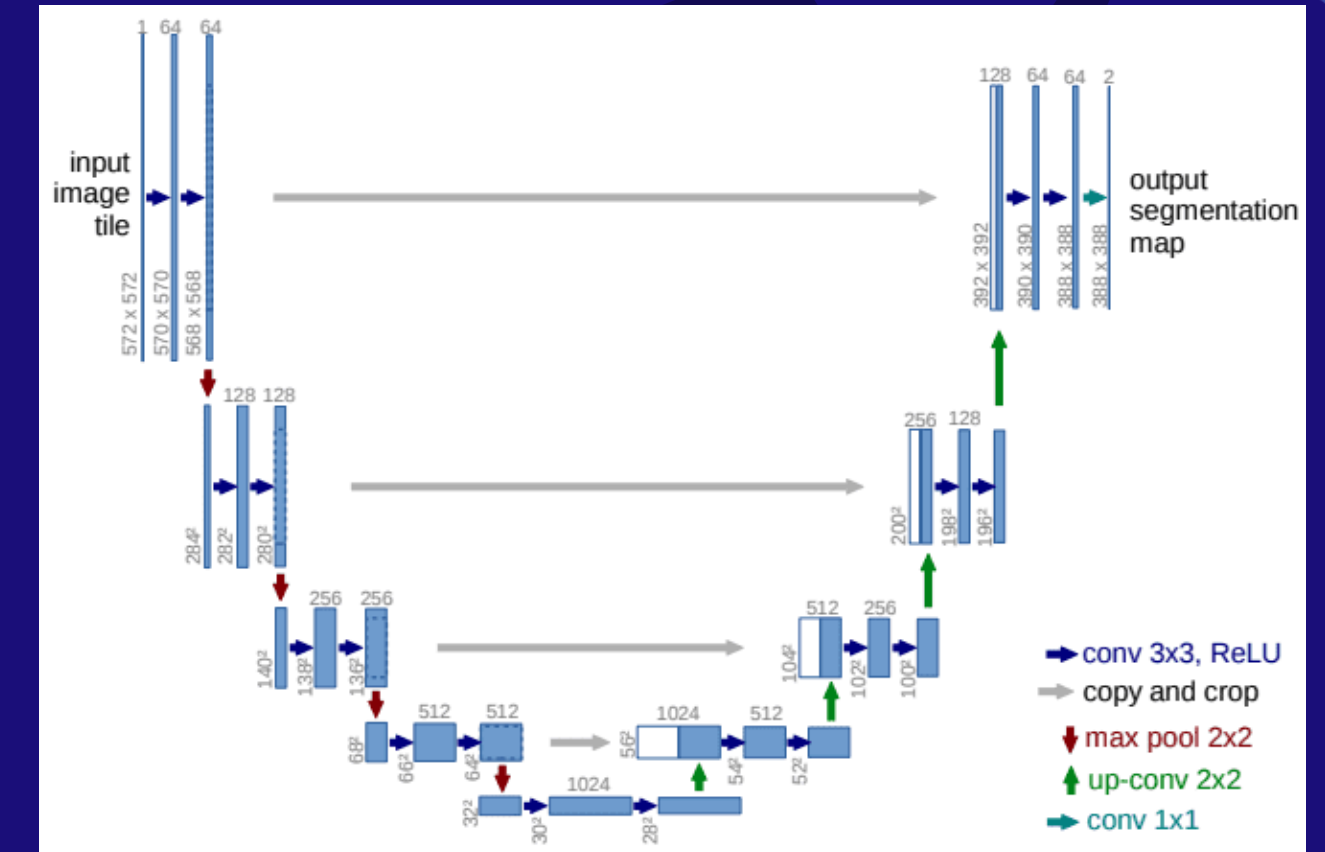
¿Por qué U-NET?

U-Net es ideal para segmentación semántica porque combina contexto global y localización precisa, produce máscaras precisas pixel a pixel, y funciona bien incluso con datasets pequeños y limitaciones de cómputo.

EFFICIENTNET-B2

Encoder

EfficientNet-B2 no usa capas MaxPool clásicas, en su lugar, hace downsampling usando convoluciones con un stride 1 dentro de los bloques MBConv. tamaño de kernel de convolución puede variar entre 3x3 o 5x5 y no tiene padding



(2025, Pepe Cantoral Ph.D). Youtube.com. https://www.youtube.com/watch?v=x_cY9l1cwj0
(2019,Mingxing Tan).EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.
<https://proceedings.mlr.press/v97/tan19a/tan19a.pdf>

Entrenamiento

FUNCIÓN DE PÉRDIDA

Cross-Entropy Loss:

$$L = - \sum_{i=1}^K y_i \log(\hat{y}_i)$$

Acelerador de Kaggle en
GPU-P100

max_lr = 1e-3
epoch = 10
weight_decay = 1e-4
mini-batch = 24 imagenes

MÉTRICAS UTILIZADAS

Accuracy, mIoU :

El mIoU mide la superposición entre las áreas predichas y reales para la segmentación.

OPTIMIZADOR

$$\lambda \sum_{i=1}^n w_i^2$$

AdamW:

Combina el momentum con la regularization de L2 para optimizar el learning rate.

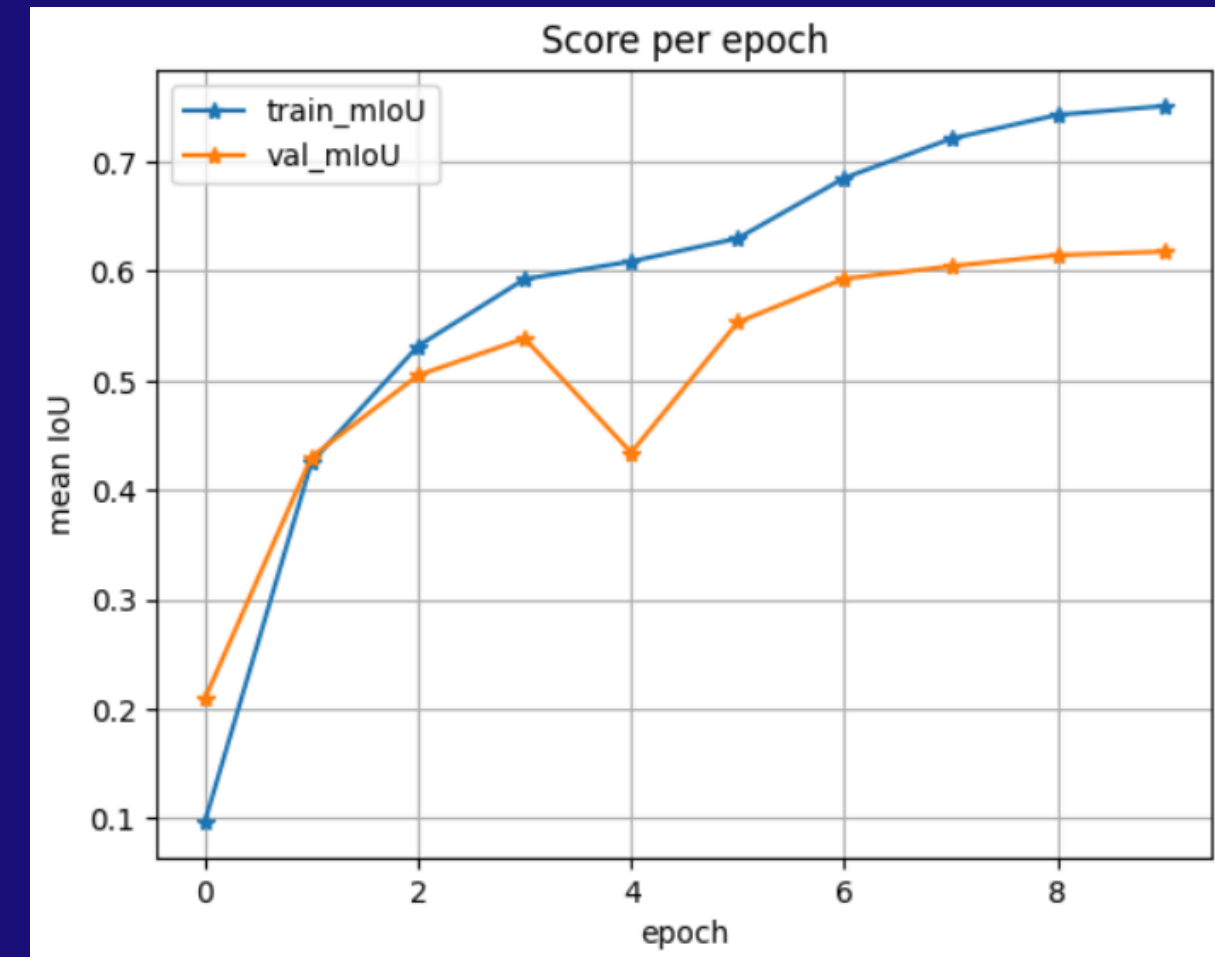
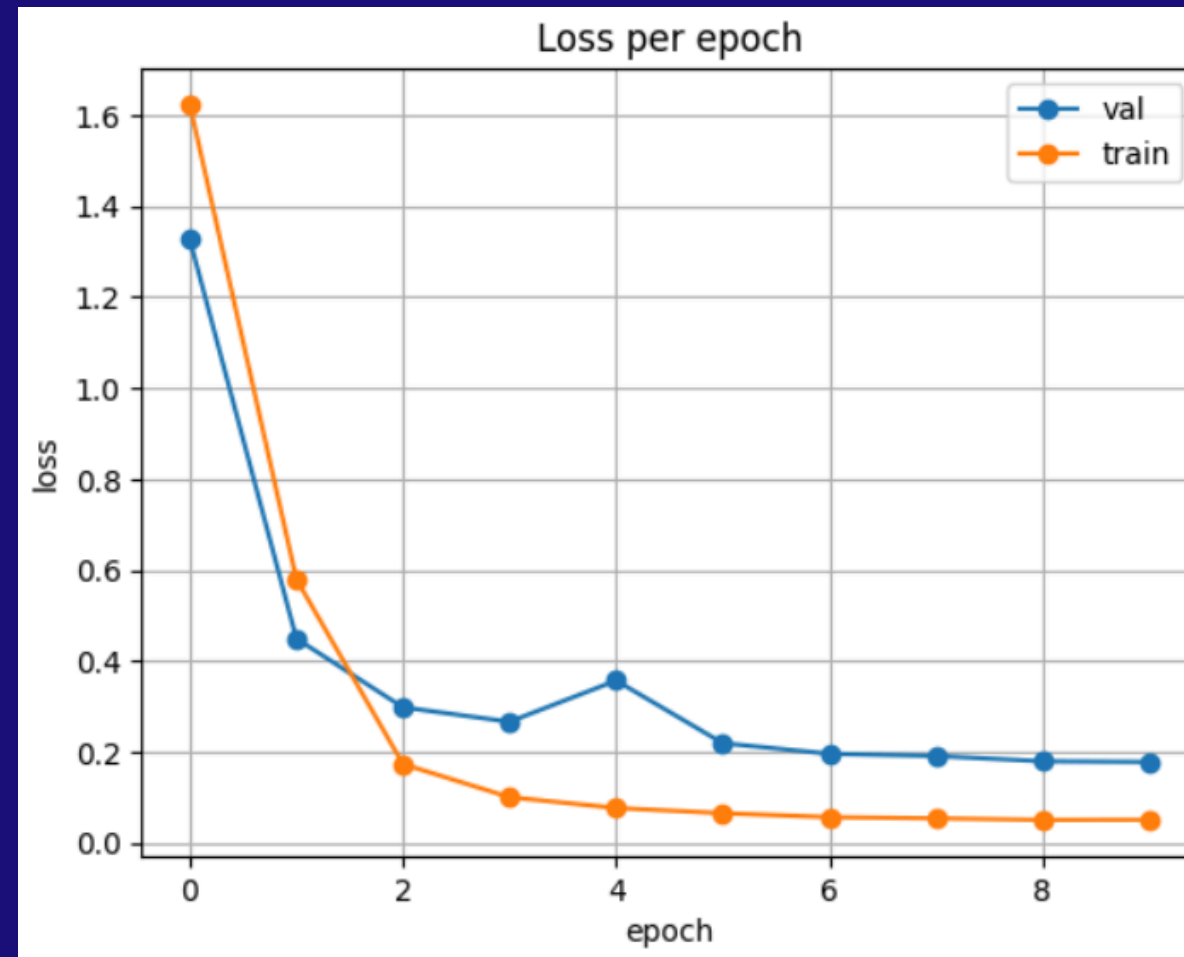
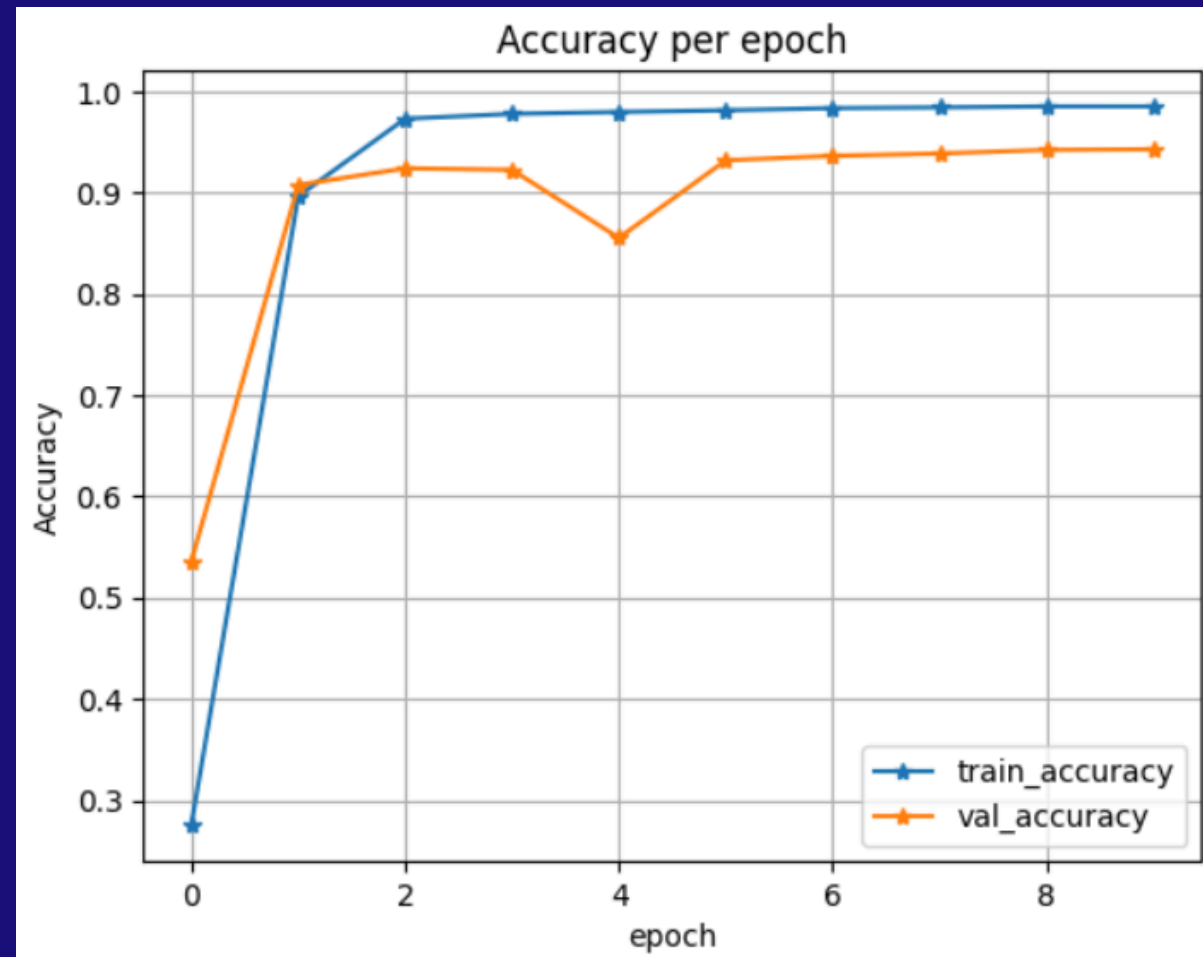
TÉCNICAS DE REGULARIZACIÓN

- Weight decay
- Early stopping

LEARNING SCHEDULE

OneCycleLearningRate: Consiste en aumentar gradualmente la tasa de aprendizaje hasta un valor máximo y luego disminuirla gradualmente hasta un valor más pequeño.

Entrenamiento



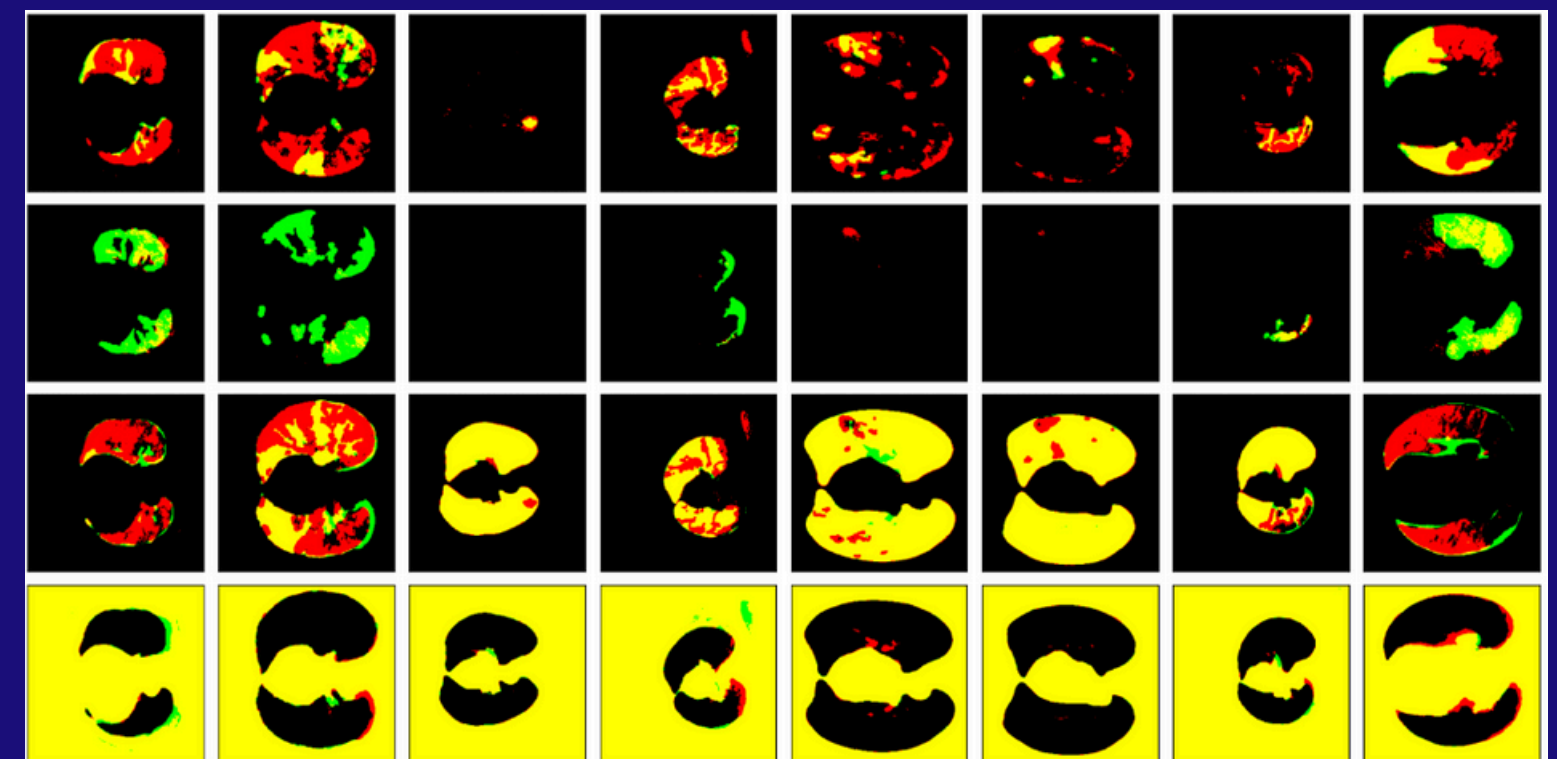
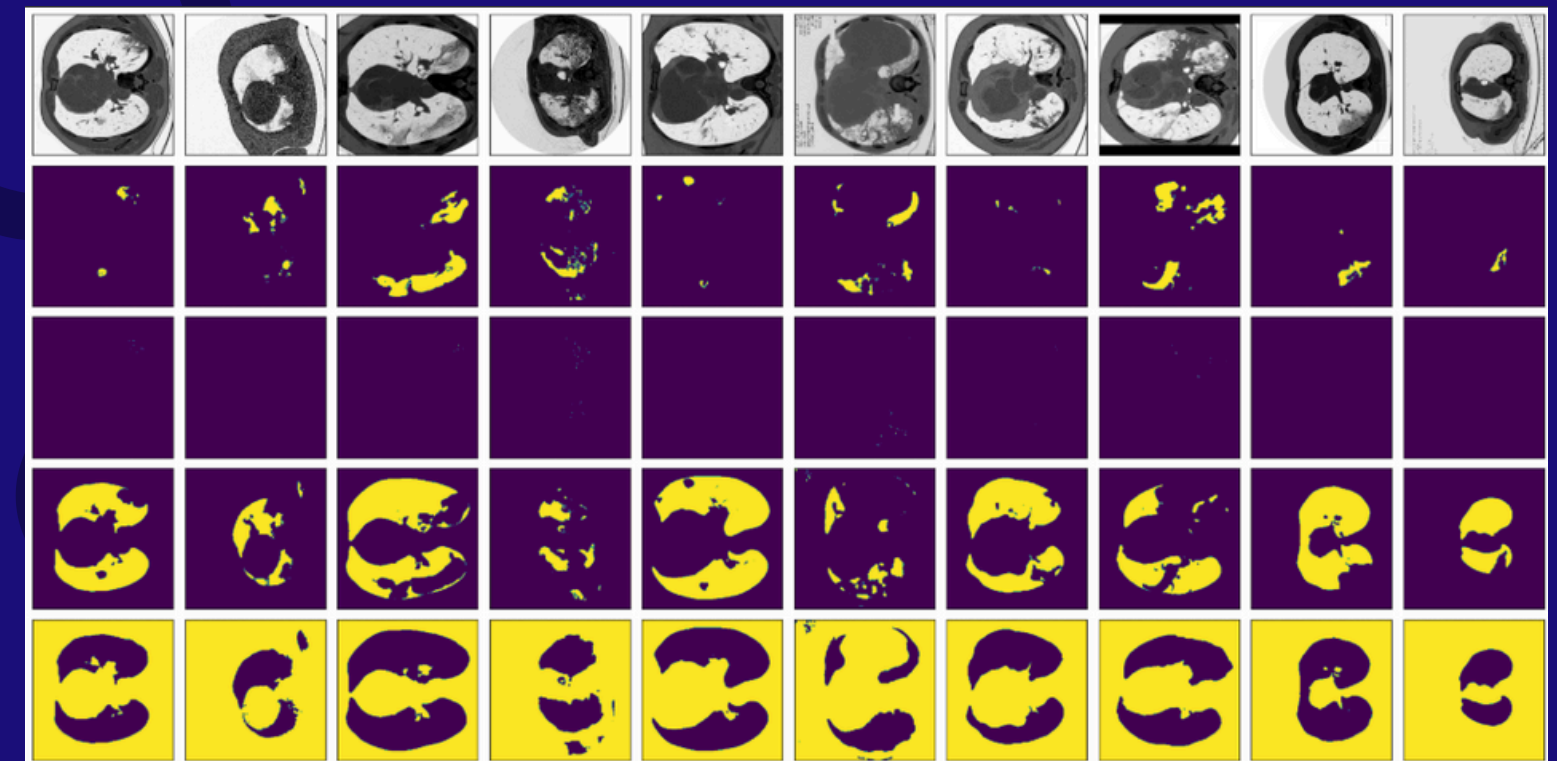
Resultados

Fila 1 = imágenes con predicciones superpuestas.

Fila 2 = etiquetas verdaderas (radiólogo).

Fila 3 = comparación predicciones vs etiquetas.

Fila 4 = salida “cruda” de tu red (qué clase asignó a cada píxel).



Conclusiones

Fortalezas

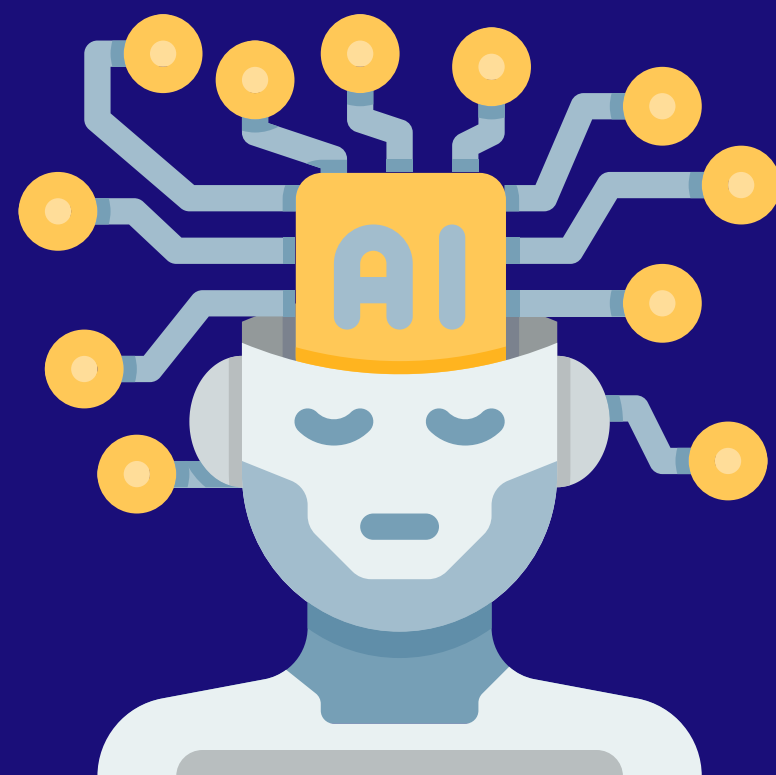
- El modelo logra un mIoU aceptable (≈ 0.57), lo que indica que ya es capaz de identificar regiones de interés (vidrio deslustrado y consolidaciones) de forma consistente.
- La arquitectura U-Net con backbones de `segmentation_models_pytorch` es adecuada para imágenes médicas, porque aprovecha bien los patrones espaciales.
- El preprocesamiento aplicado (clipping, normalización y aumento de datos con rotaciones y crops) aporta robustez frente a la variabilidad de los cortes de tomografía.

Posibles Mejoras

- Incrementar la calidad y cantidad de datos (más volúmenes 3D o usar conjuntos externos como MosMed).
- Mejorar la consistencia de anotaciones para evitar solapamientos entre clases en las máscaras.
- Explorar arquitecturas más avanzadas (ej. U-Net++ o modelos con atención) que mejoren la detección en áreas difíciles.
- Implementar métricas adicionales (ej. Dice coefficient) para complementar la evaluación y tener una visión más clínica.

Debilidades

- El $mIoU < 0.6$ revela que aún existe un desajuste importante respecto a las segmentaciones de referencia, especialmente en los bordes o en casos de lesiones pequeñas.
- El dataset de entrenamiento es relativamente pequeño (≈ 900 imágenes) y contiene alta variabilidad entre pacientes, lo que limita la generalización del modelo.
- Posible confusión entre clases (ej. ground-glass vs consolidación) reflejada en las zonas donde se mezclan colores, lo que baja la precisión en lesiones complejas.
- El criterio de early stopping basado solo en pérdida puede detener el entrenamiento antes de que el modelo refine mejor las fronteras.



Gracias

Referencias

Cantoral Ph.D (2025). Youtube.com. https://www.youtube.com/watch?v=x_cY9l1cwj0

Cantoral Ph.D (2025). Youtube.com. <https://www.youtube.com/watch?v=dfMEK4bKjRE>

Igor.Slinko. COVID-19 CT Images Segmentation . <https://kaggle.com/competitions/covid-segmentation>, 2020. Kaggle.

Maede Maftouni. (2021, May 8). PyTorch Baseline for Semantic Segmentation. Kaggle.com; Kaggle. <https://www.kaggle.com/code/maedemaftouni/pytorch-baseline-for-semantic-segmentation>

Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. ArXiv.org. <https://arxiv.org/abs/1905.11946>

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. ArXiv.org. <https://arxiv.org/abs/1505.04597>