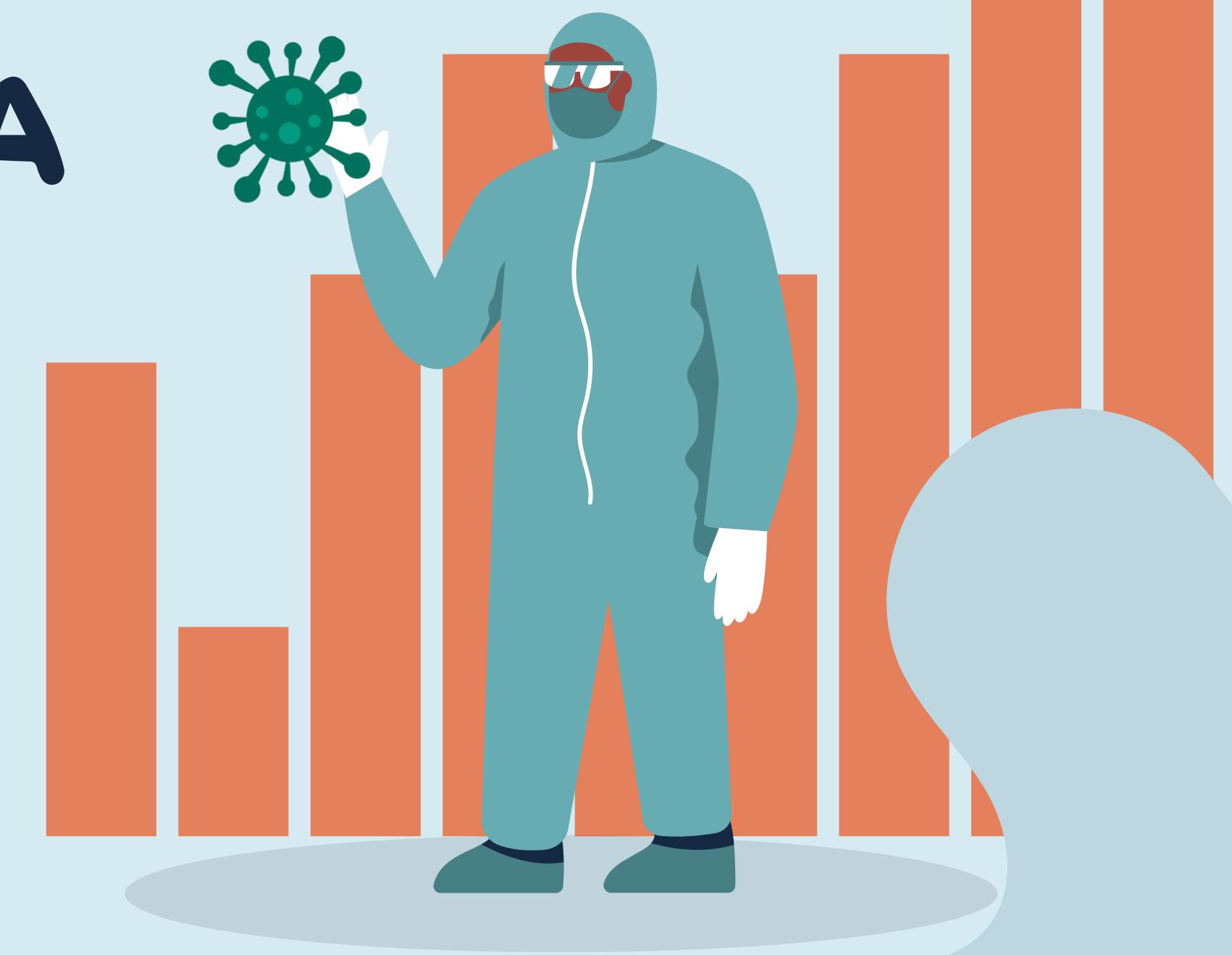


SEGMENTACIÓN DE IMAGENES PARA LA DETECCIÓN DEL COVID-19

Labubus, Diddys, Nitos, Lattes, Matchas y Lululemons 



TEAM MEMBERS



Angel Ramirez
A01710158



Diego Garcia
A01710777



Kevin Ramirez
A01711063



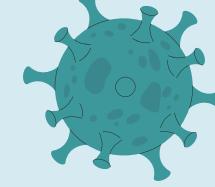
Eduardo Viveros
A01710605



Paulina López
A01701095



Cristian Chavez
A01710680

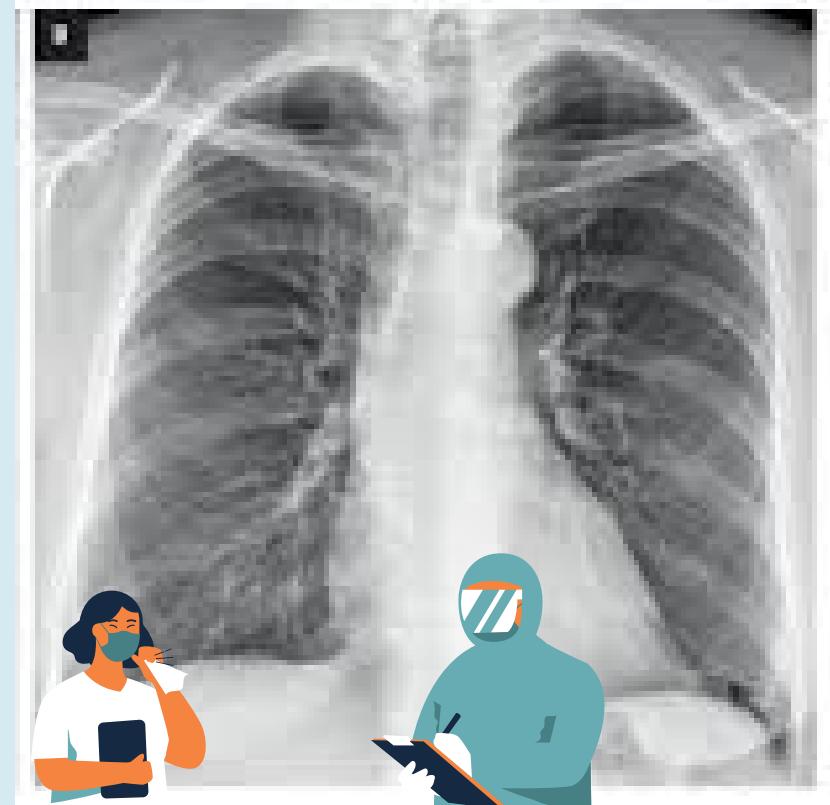


SITUACIÓN PROBLEMA - SEGMENTACIÓN MEDICA PARA COVID-19

El COVID-19 ingresa al cuerpo principalmente por vías respiratorias dañando a las células alveolares, reduciendo la capacidad de oxigenación en la sangre.

Las imágenes CT (tomografías computarizadas) muestran patrones característicos:

- Opacidades en vidrio esmerilado (Ground-glass)
 - Hallazgo temprano: Áreas grises y difusas, inflamación leve o acumulación de líquido en alvéolos.
- Consolidaciones
 - Enfermedad más avanzada o grave: Áreas blancas y densas, acumulación de líquido y células inflamatorias.



Importancia de la segmentación en CT:

- Cuantificar la extensión del daño pulmonar.
- Monitorear la progresión de la enfermedad.
- Evaluar la respuesta al tratamiento.
- Asistir en el diagnóstico rápido.



- Objetivo: Desarrollar modelos de segmentación médica para detectar zonas infectadas en pulmones a partir de tomografías.
- Datos disponibles:
 - * Imágenes de tomografía computarizada (CT).
 - * Máscaras anotadas por especialistas que delimitan áreas infectadas.
- Propósito: Mejorar la precisión y velocidad del diagnóstico, apoyando a radiólogos en la lucha contra la pandemia.

- Impacto clínico: Apoya a radiólogos en diagnósticos rápidos y precisos.
- Reduce variabilidad entre especialistas y tiempos de análisis.

IMPLICACIONES ETICAS

Las implicaciones éticas en la segmentación de imágenes CT médicas.

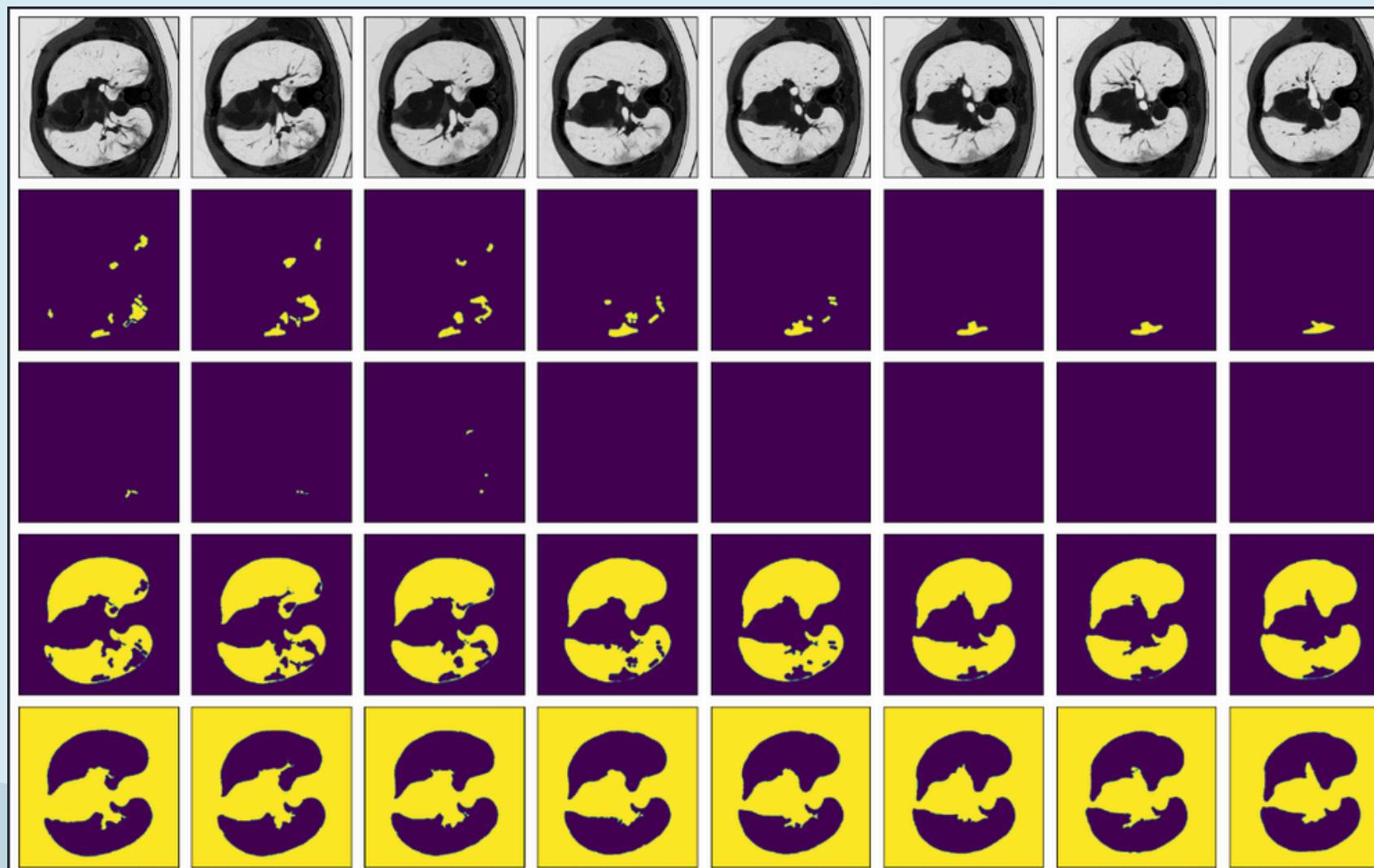
- Privacidad y Confidencialidad:
 - Las imágenes contienen información sensible del paciente por lo que es crucial el anonimato de los datos, eliminando metadatos identificables, además de obtener el consentimiento informado para su uso en investigación cumpliendo regularizaciones:
 - GDPR (Europa)
 - HIPPA (EE.UU)
- Sesgo en datos:
 - Es esencial usar datasets diversos y representativos.
- Uso clínico vs experimental
 - Los modelos en fase de investigación no deben usarse clínicamente sin validación rigurosa.
Aprobación ética y ensayos clínicos para garantizar seguridad y eficacia.



ANALISIS DE DATOS



PROCESO ETL/EDA



Donde, cada canal representa una clase distinta:

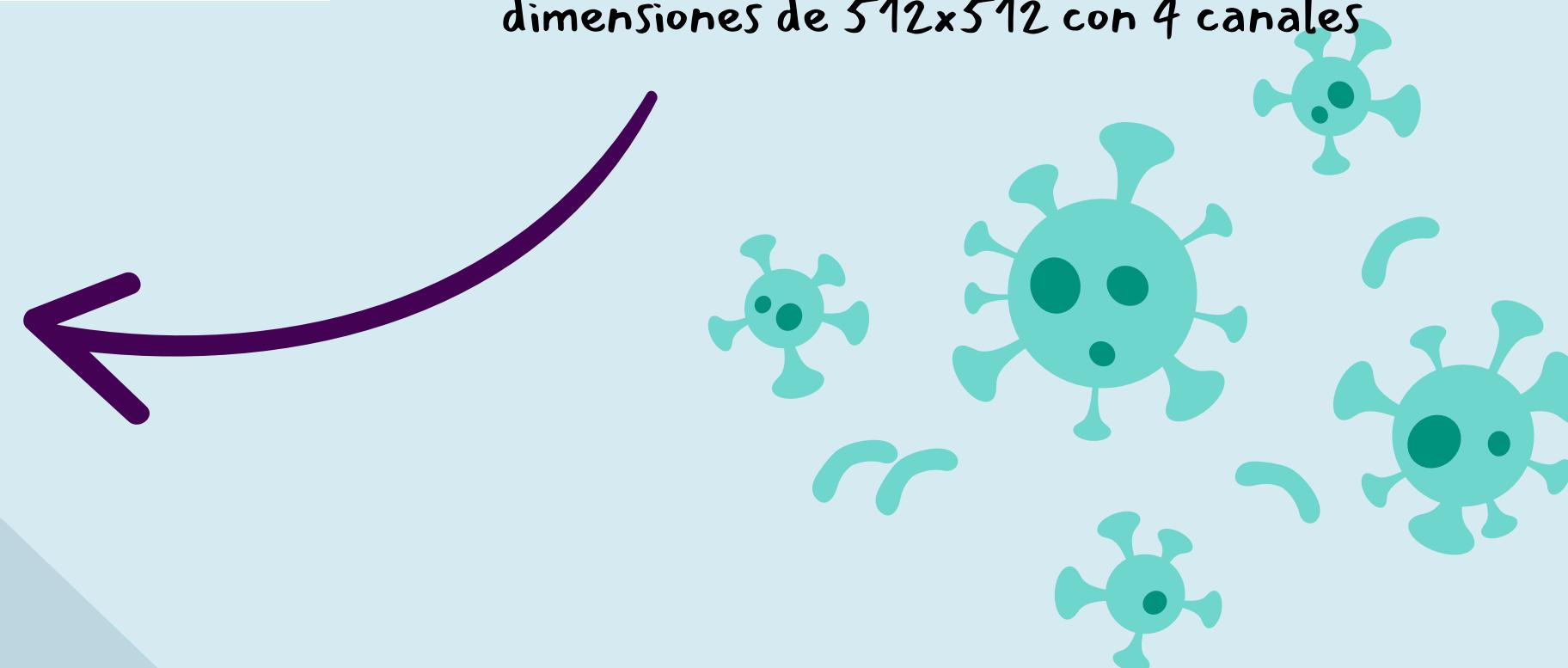
- Class [0] - "Ground glass" - Target
- Class [1] - "Consolidations" - Target
- Class [2] - "Lungs other" - No importante
- Class [3] - "Background" - No importante

El dataset proporcionado, incluye tomografías computarizadas (CT) del pulmón con las máscaras de segmentación correspondientes.

Las imágenes TC proceden de dos conjuntos de datos:

- Radiopedia - 829 imágenes
- Medseg - 100 imágenes + 24 de test

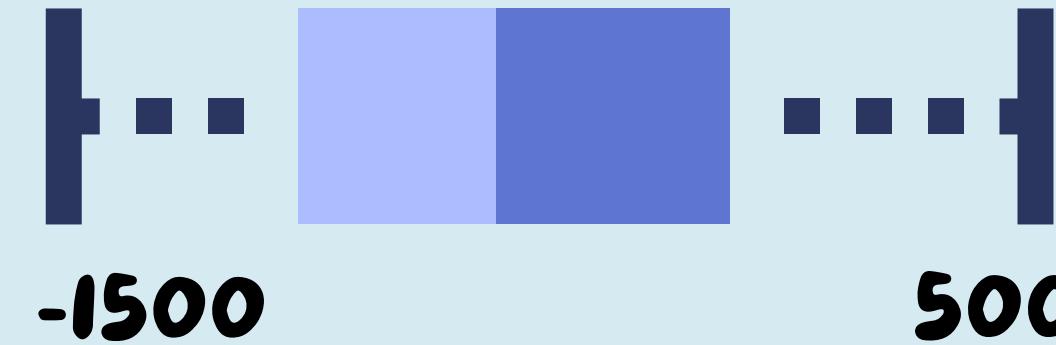
Después del procesamiento, todas las imágenes tienen las dimensiones de 512x512 con 4 canales



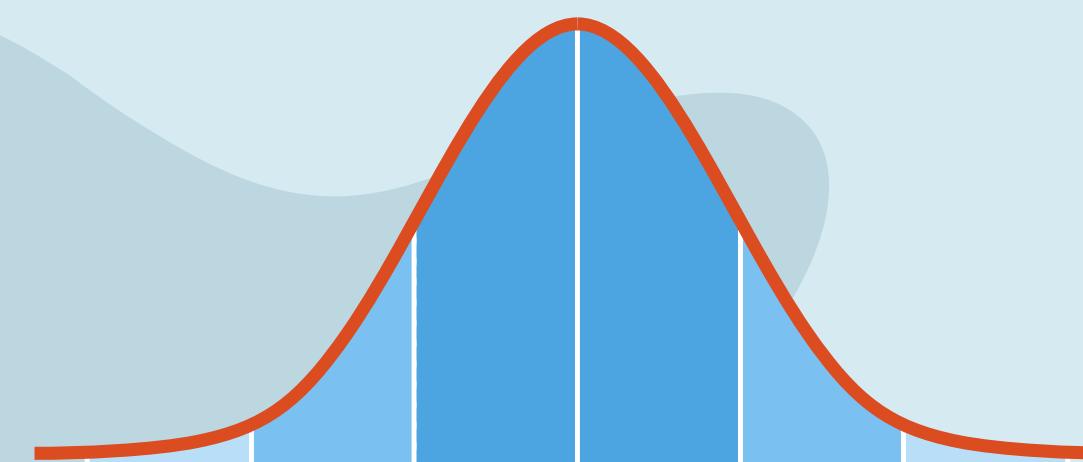
PREPROCESAMIENTO DE DATOS I

Preprocesamiento de datos

- Clipping y uso de percentiles para valores mayores a >500 y menores a <-1500 como filtro para la eliminación de pixeles (medidos en unidades Hounsfield ~ diferentes niveles de radiodensidad de los tejidos humanos.)



- Estandarización para features con un mismo(o similar) rango



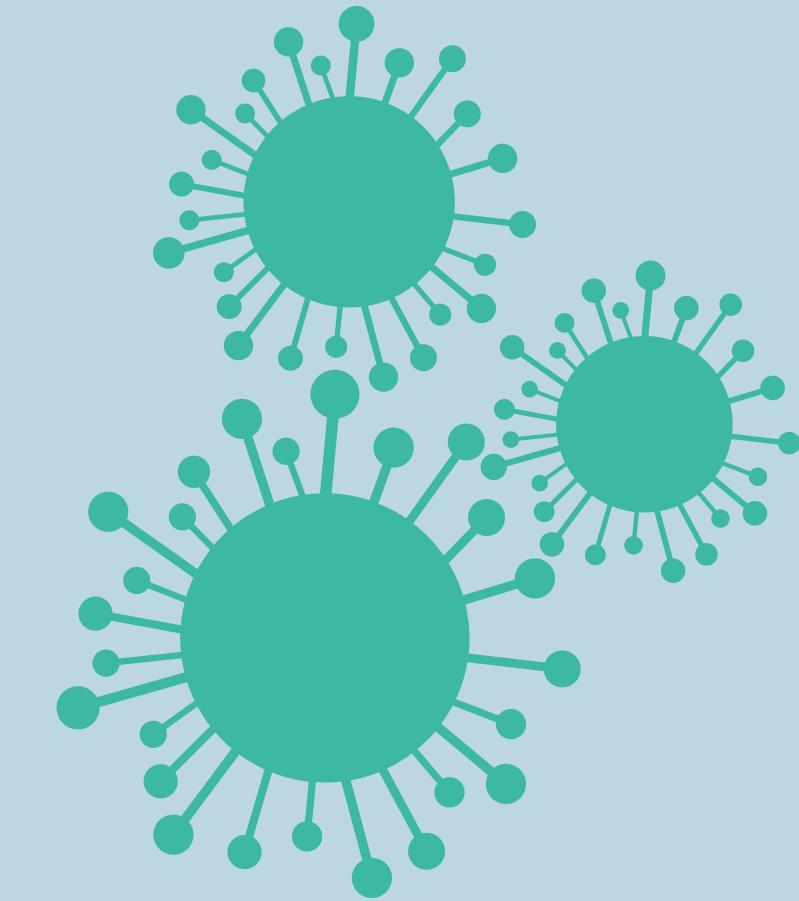
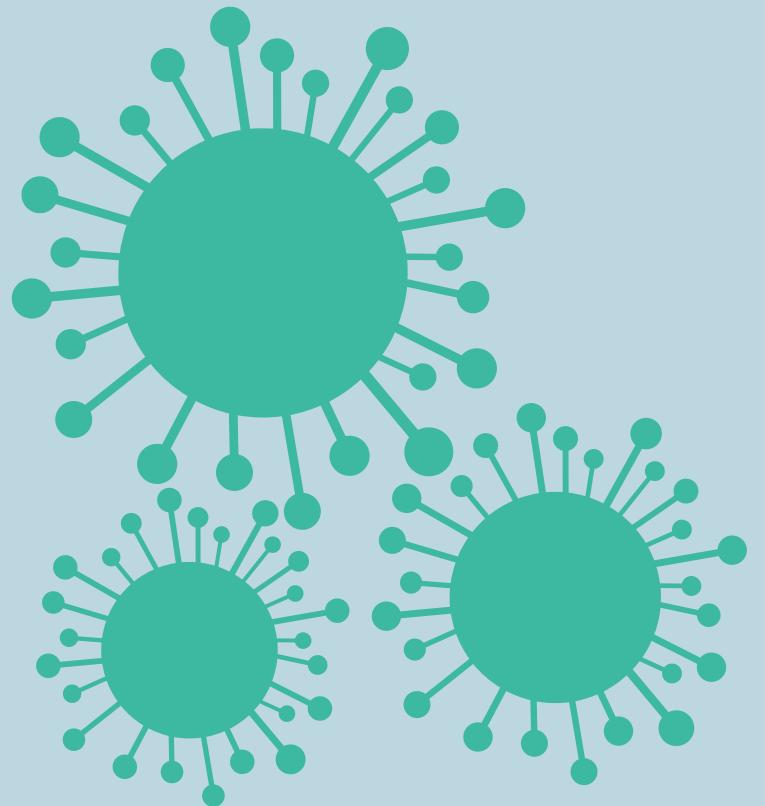
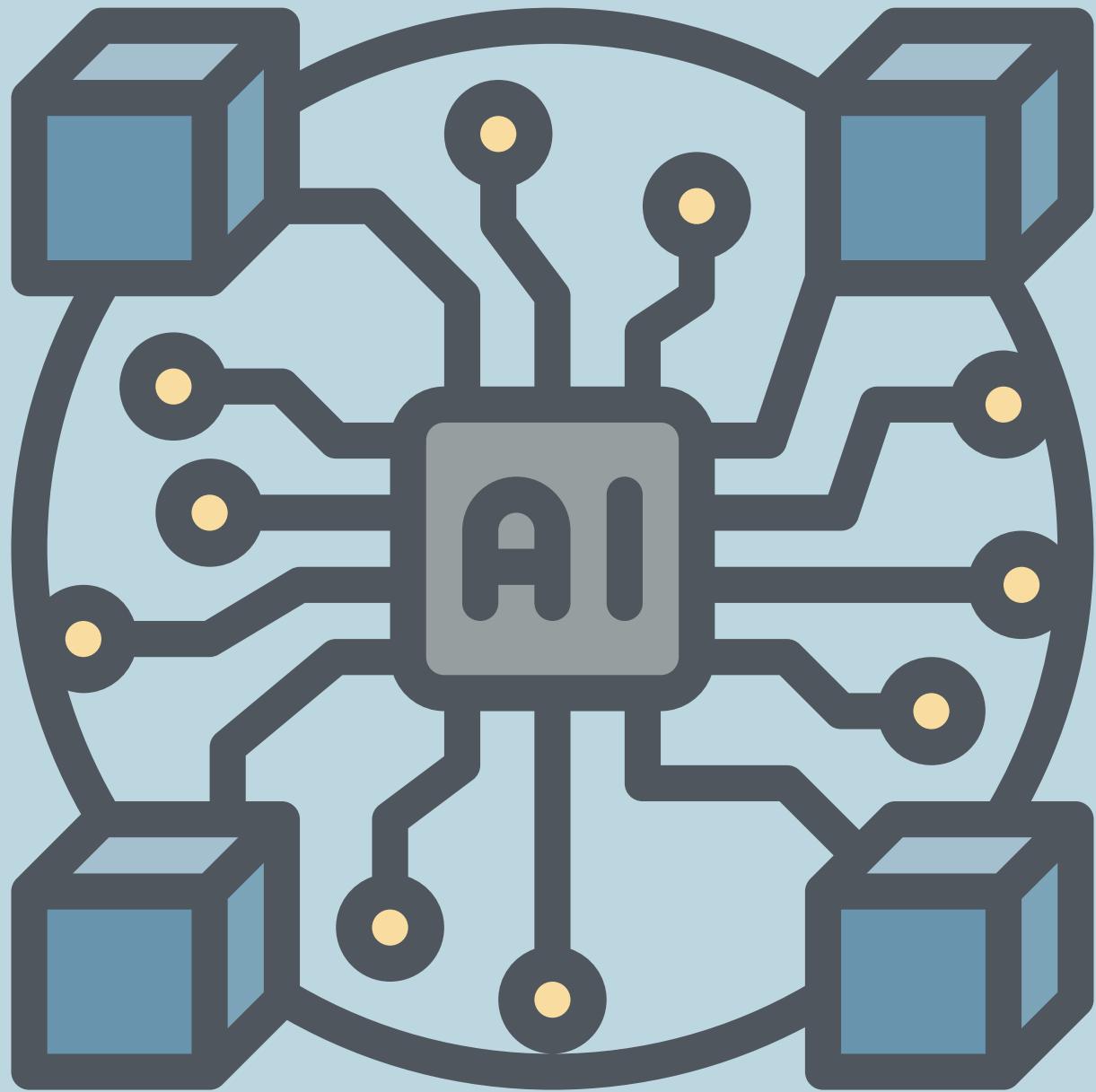
TEJIDO	HU
Hueso/Calcificación	80 a 1000
Sangre coagulada	56 a 76
Sangre no coagulada	40
Sustancia gris	36 a 46
Sustancia blanca	22 a 32
Musculo	10 a 40
Líquido cefalorraquídeo	15
Grasa	-100 a -50
Aire	-1000

Data Augmentation:

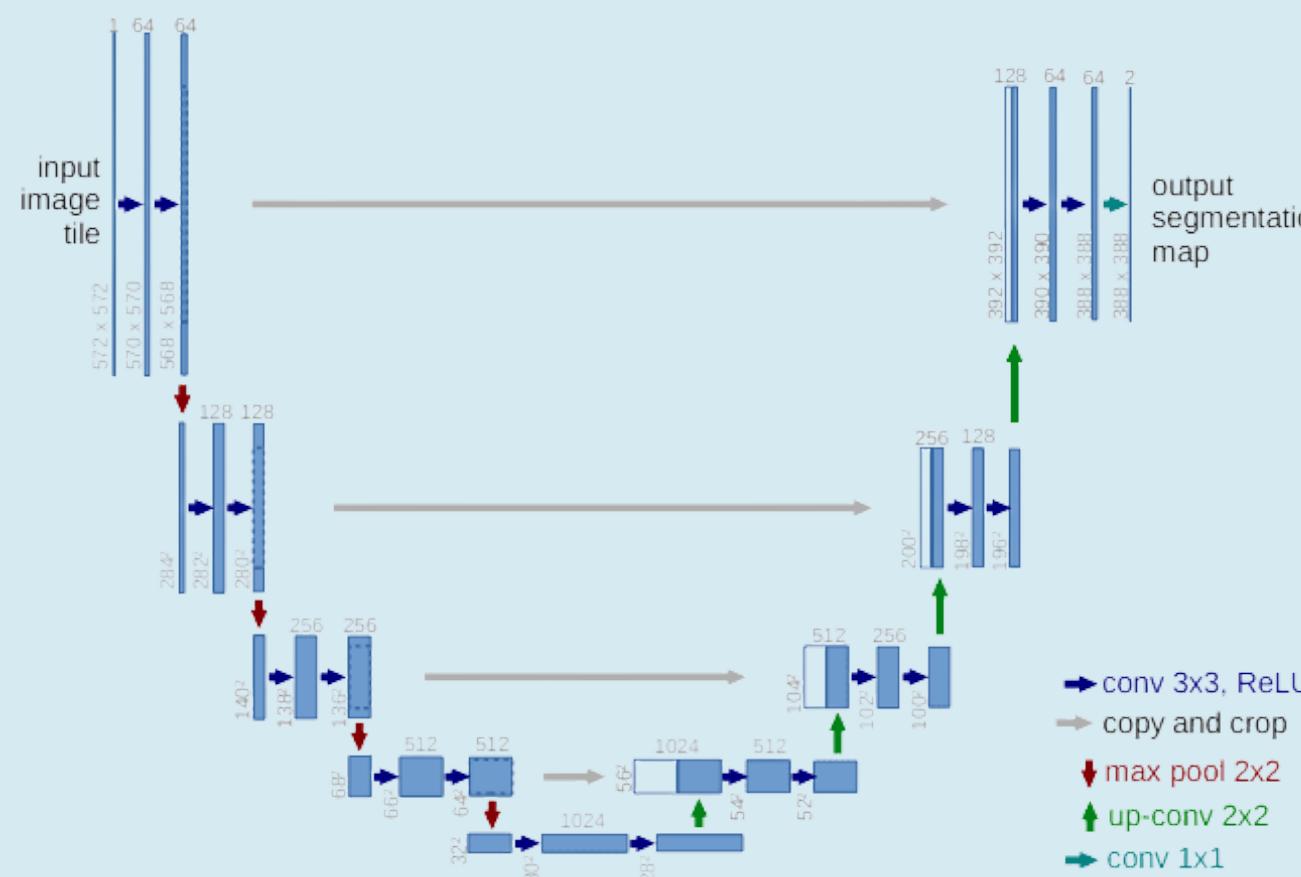
Creamos datos artificiales para el train. Donde:

- Rotación: Hasta 360 grados
- Random Crop: Redimensionamiento aleatorio
- Flip horizontal: Volteo horizontal

MODELO

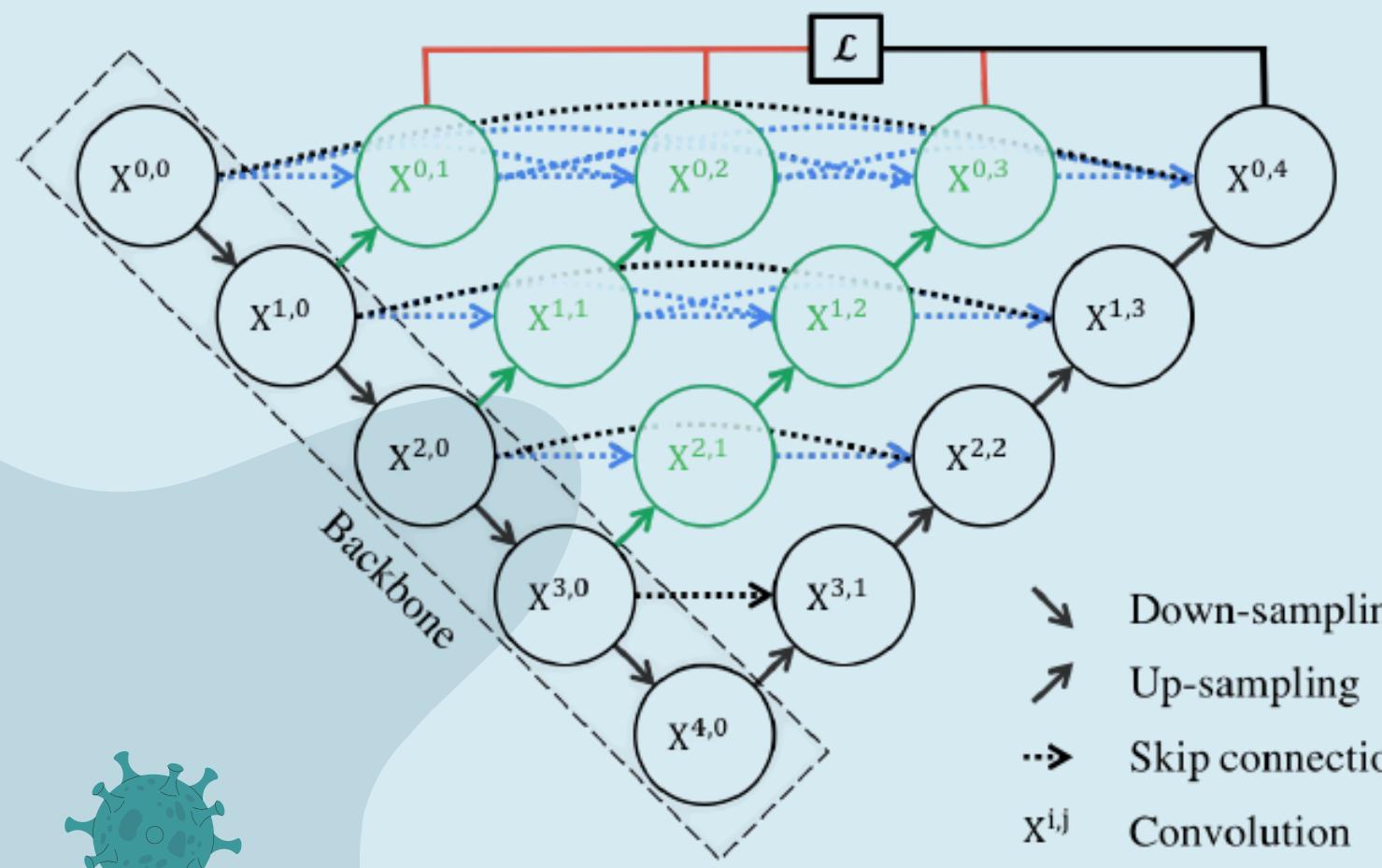


ARQUITECTURA DEL MODELO



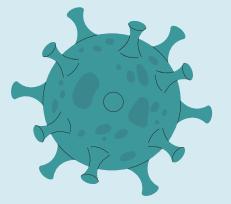
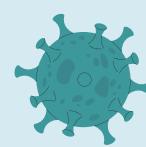
Para el entrenamiento de segmentación se utilizará la arquitectura:

- EfficientNet-b2 encoder
- U-net++ decoder
- Optimizador Adam
- Función de activación ReLU



Skip Connections

- La concatenación no es sumar pixeles.
- Es poner los mapas de características como "capas de colores diferentes".
- El ancho y alto ($H \times W$) siempre deben coincidir → por eso se usa upsampling para igualarlos.
- Lo que cambia es el número de canales (se suman todos los que concatenas).



MASK_TO_ONEHOT

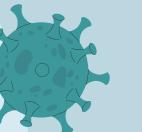
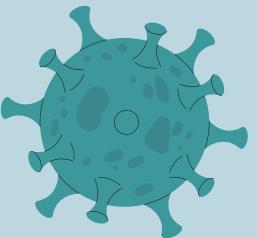
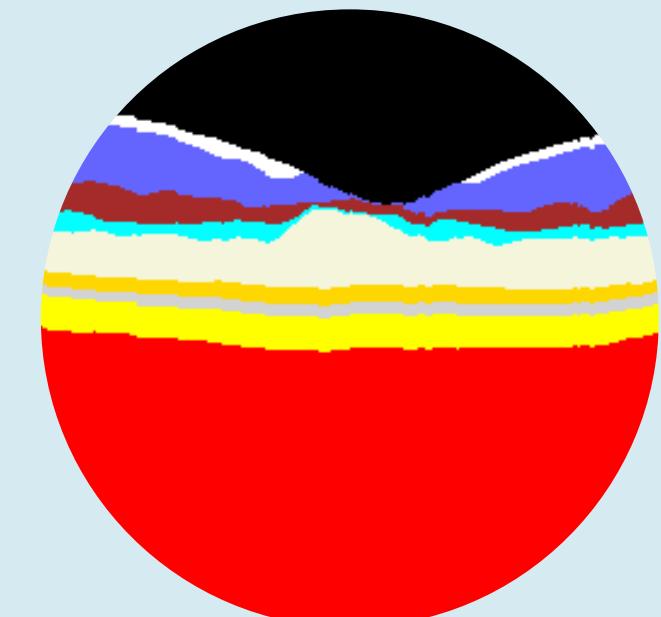
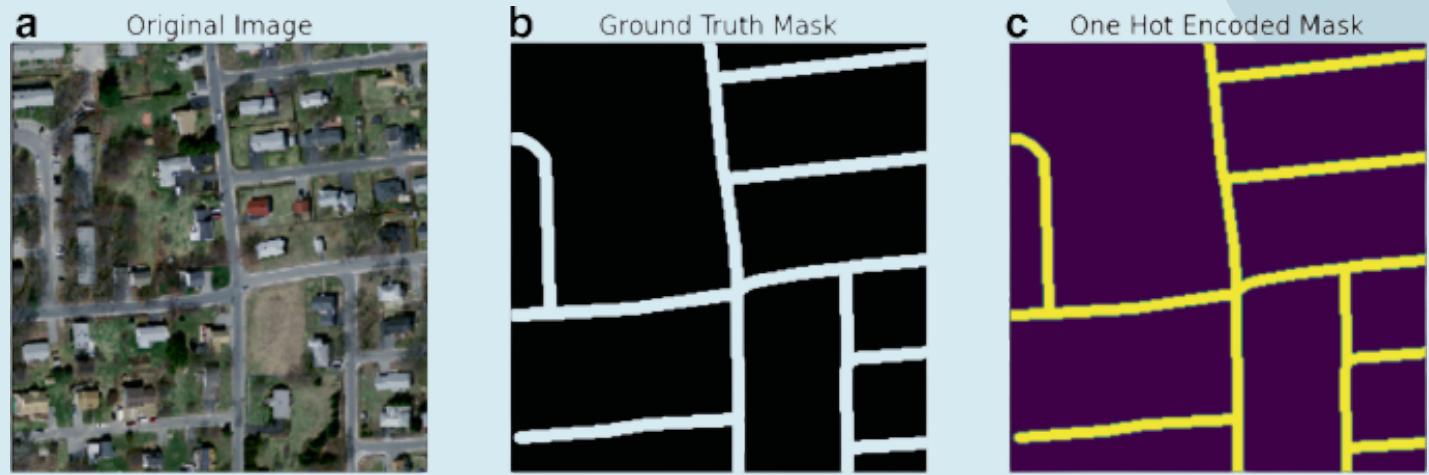
Qué hace:

- Convierte la máscara original de segmentación en 4 canales separados, uno por cada clase.
- Cada canal tiene 1 donde está esa clase y 0 en el resto.

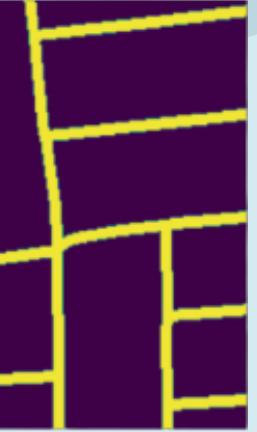
Ejemplo: si un píxel pertenece a la Clase 2 → en el canal 2 será 1 y en los demás 0.

Para qué sirve aquí:

- Visualización: puedes mostrar cada clase por separado en la imagen.
- Facilita ver claramente dónde aparece cada clase sin confundirlas con otras.



One Hot Encoded Mask



IoU (INTERSECTION OVER UNION)

Qué mide: Superposición entre la máscara predicha y la real para una clase específica.

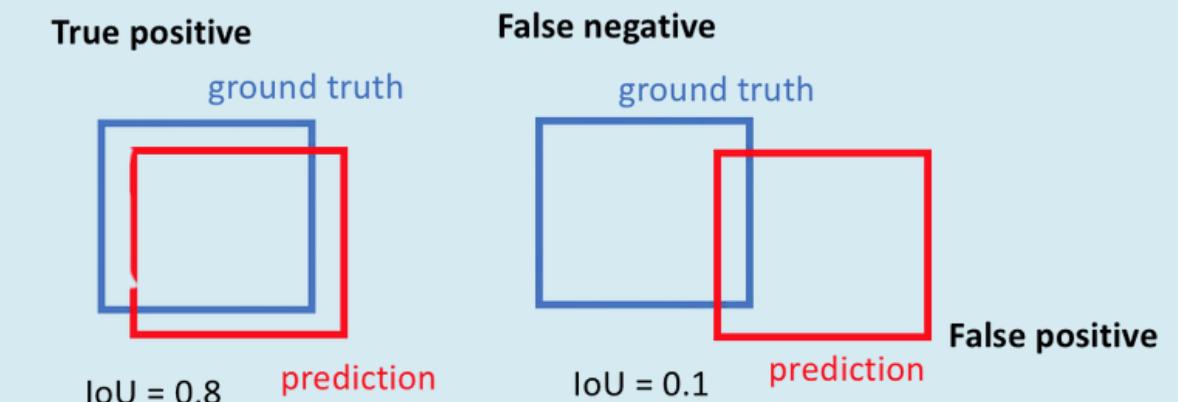
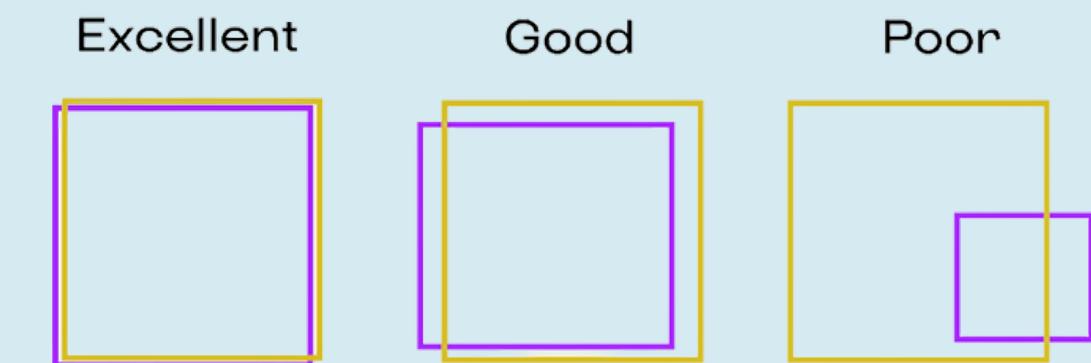
Cómo se calcula:

- Preparar máscaras → limpiar y normalizar datos.
- Etiquetar píxeles → asignar cada píxel a una clase.
- Comparar → matriz que indica coincidencias y diferencias.
- Intersección (i) → píxeles predichos correctamente.
- Unión (u) → píxeles donde aparece la clase en pred o real.
- $IoU = i / u$ (si no hay píxeles de la clase → $IoU = 1$).

Interpretación:

- IoU alto = buena predicción de esa clase.
- IoU bajo = mala segmentación o confusión.

$$IoU = \frac{|P \cap G|}{|P \cup G|}$$

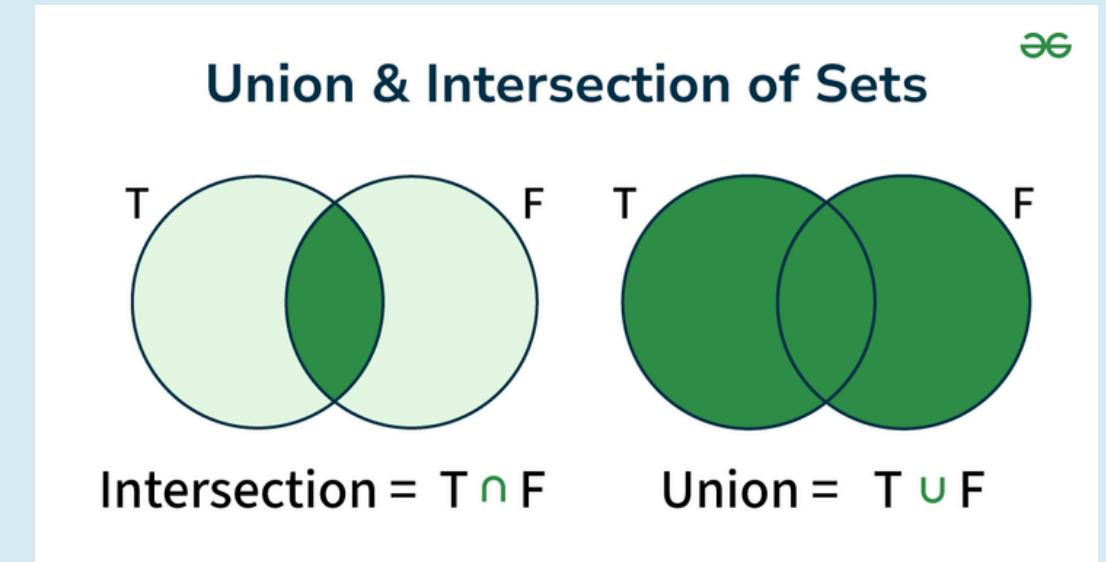


MIoU (MEAN INTERSECTION OVER UNION)

- Qué es: Métrica clave en segmentación de imágenes. Evalúa el traslape entre la máscara predicha y la real, por clase.

Proceso de cálculo:

- Softmax: Se generan probabilidades por clase para cada píxel.
- Predicción: Se asigna a cada píxel la clase con mayor probabilidad.
- Aplanar: Predicción y máscara real se convierten en vectores para comparar píxel a píxel.
- IoU por clase:
 - Intersección = píxeles predichos correctamente.
 - Unión = píxeles presentes en al menos una de las dos máscaras.
 - Clases ausentes en la real se ignoran.
- Promedio: Se promedian los IoU de todas las clases \rightarrow mIoU final.
- Interpretación: Un mayor mIoU indica que el modelo segmenta con mayor precisión.



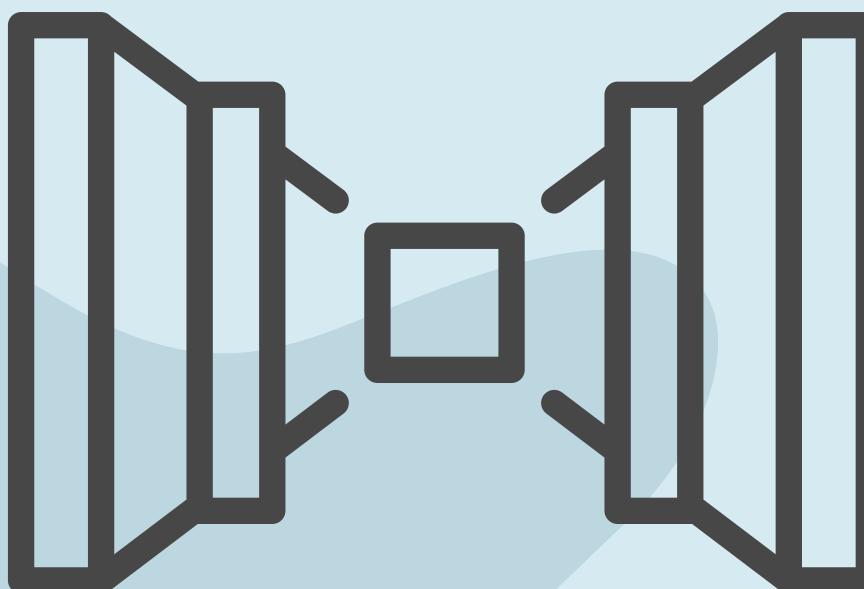
EFFICIENTNET-B2 (ENCODER)

Funciona como "extractor de características", utiliza una técnica llamada "escalado compuesto" que ajusta simultáneamente:

- Profundidad (número de capas)
- Ancho (número de filtros)
- Resolución (tamaño de entrada)

Características:

- Número de capas: 16 capas convolucionales (codificador)
- Tamaño de kernel: 3×3 .
- Stride: Capas iniciales 2 para reducir rápidamente la dimensionalidad, capas posteriores 1 para mantener resolución durante extracción de características profundas.
- Padding: Usa padding "same", el tamaño de salida permanece igual al de la entrada, conserva resolución.
- Activación: ReLU



U-NET++ (DECODER)

Características:

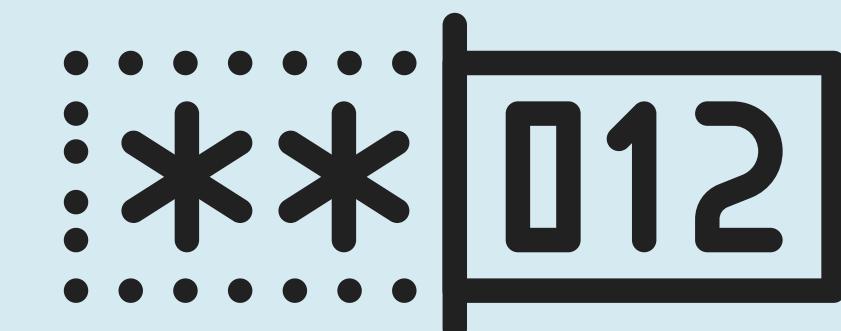
- Número de capas: 5 por rama de decoder
- Tamaño de kernel: 3×3
- Stride: 1
- Padding: Usa padding "same", el tamaño de salida permanece igual al de la entrada, conserva resolución.
- Upsampling:
 - Se usa transposed convolution, kernel 2×2 y stride de 2
- Skip Connections:
 - Pasan por bloques convolucionales adicionales para alienar características antes de fusionarlas.

Capas finales:

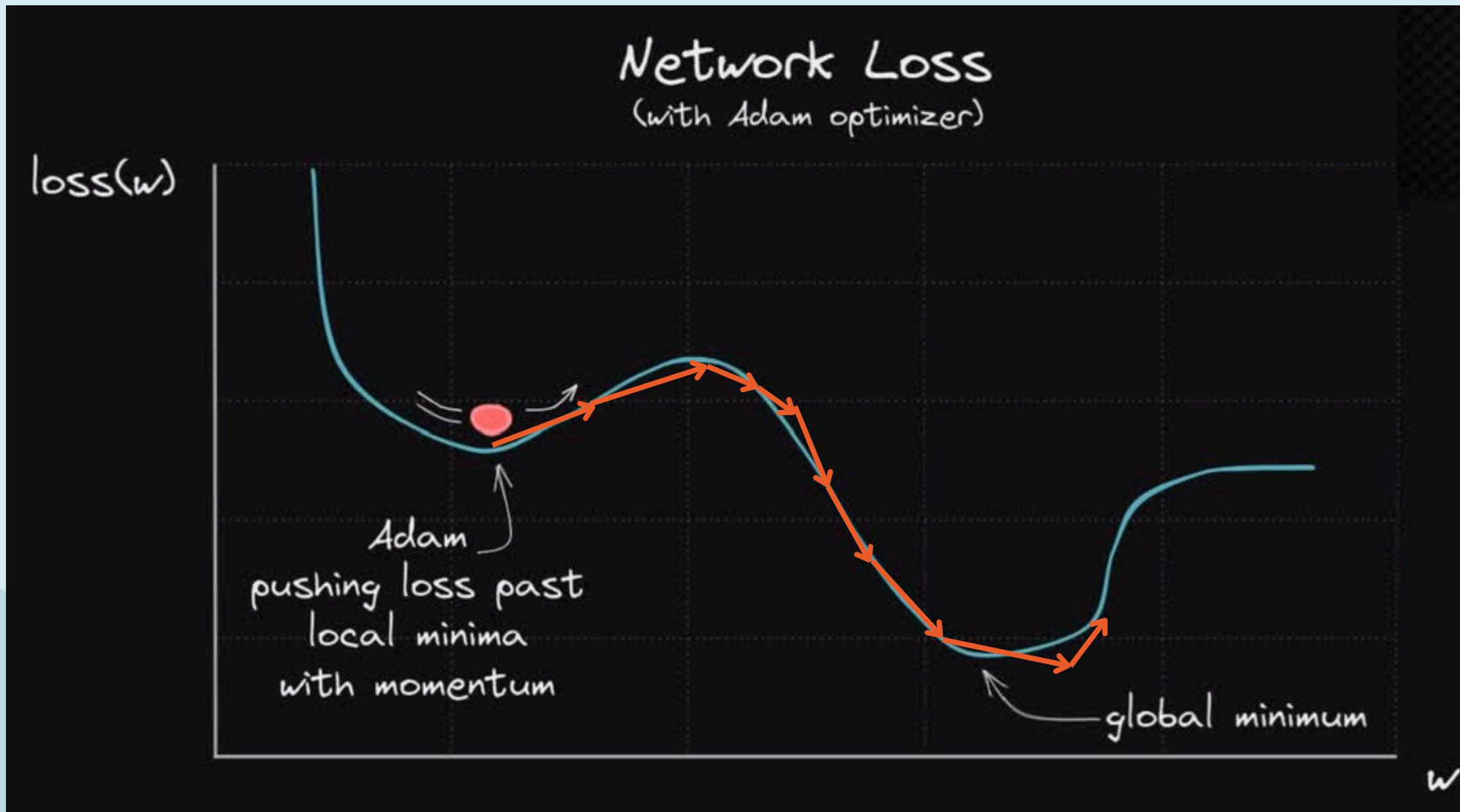
- Activación:
 - Softmax, multiclasificación exclusiva

Función de pérdida:

- Cross-Entropy Loss:
 - Penaliza predicciones incorrectas por pixel



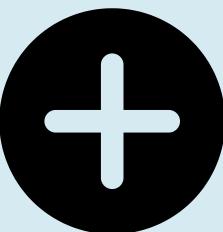
OPTIMIZADOR - ADAM



Adam es uno de los algoritmos de optimización más populares y efectivos en el campo del Deep Learning.

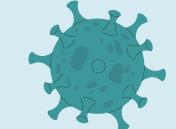


SGD &
Momentum: Se
gana impulso en
una dirección.



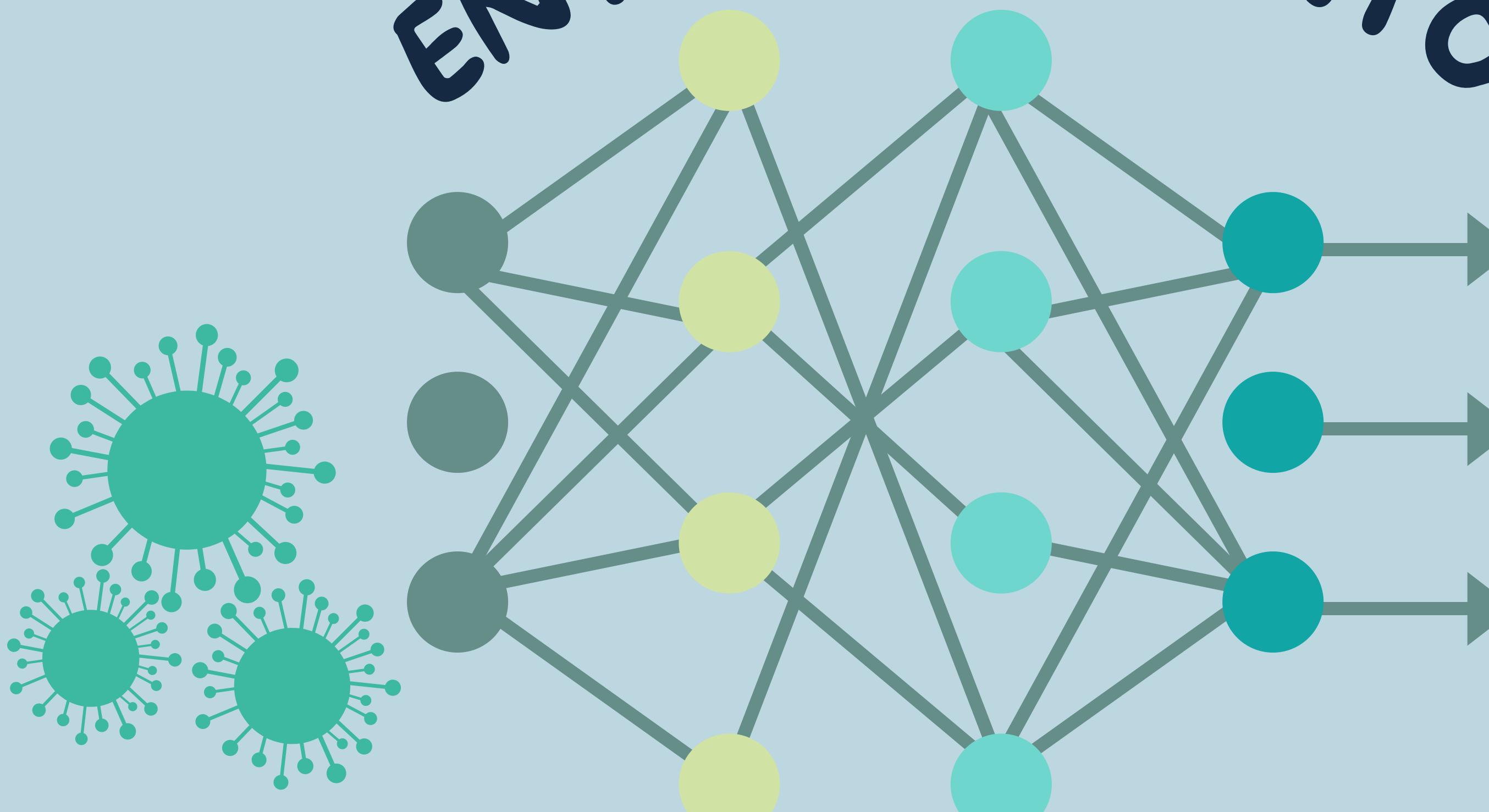
RMSProp : adaptar
el learning rate
para cada
parámetro.

α

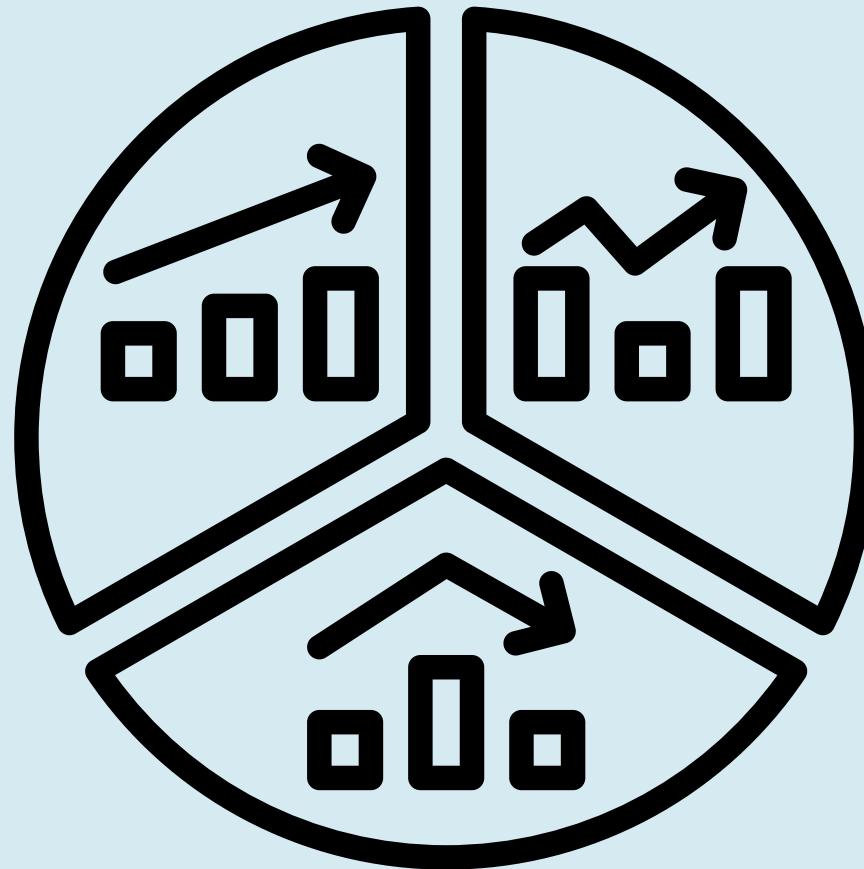


ADAM

ENTRENAMIENTO



SET DE ENTRENAMIENTO/VALIDACIÓN/PRUEBAS

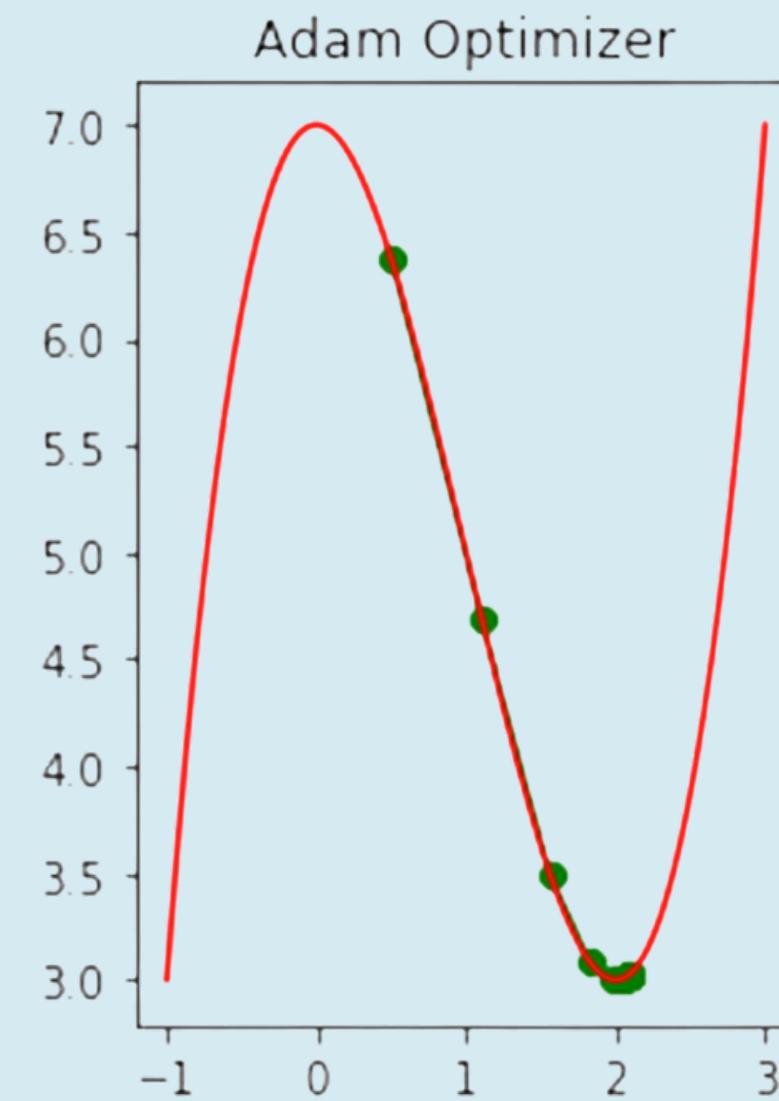


- Para validación se utilizan las primeras 24 instancias de MedSeg
 - Para entrenamiento se usa el resto de MedSeg adicionando todo Radiopedia
 - Para test solo es para inferencia y sin usar etiquetas
-
- La validación se separa desde el inicio para verificar que no haya overfitting posteriormente en train.
 - Se mezcla Radiopedia + MedSeg en train mejora la generalización del modelo de segmentación ya que son 2 dominios distintos.
 - El test permite reportar resultados sin sobreajuste a la validación.

HIPERPARAMETROS Y REGULARIZACIÓN



- Learning rate
- Numero de epochs
- Tamaño de minibatch
- Early stopping
- Early stopping:
 - Si la pérdida de val no mejora después de 7 epochs seguidas, se detiene el entrenamiento, evitando overfitting
- Guardado del mejor modelo:
 - El modelo se guarda cada que la perdida de val disminuya, con esto garantizamos guardar el mejor modelo.



PROCESO DE ENTRENAMIENTO

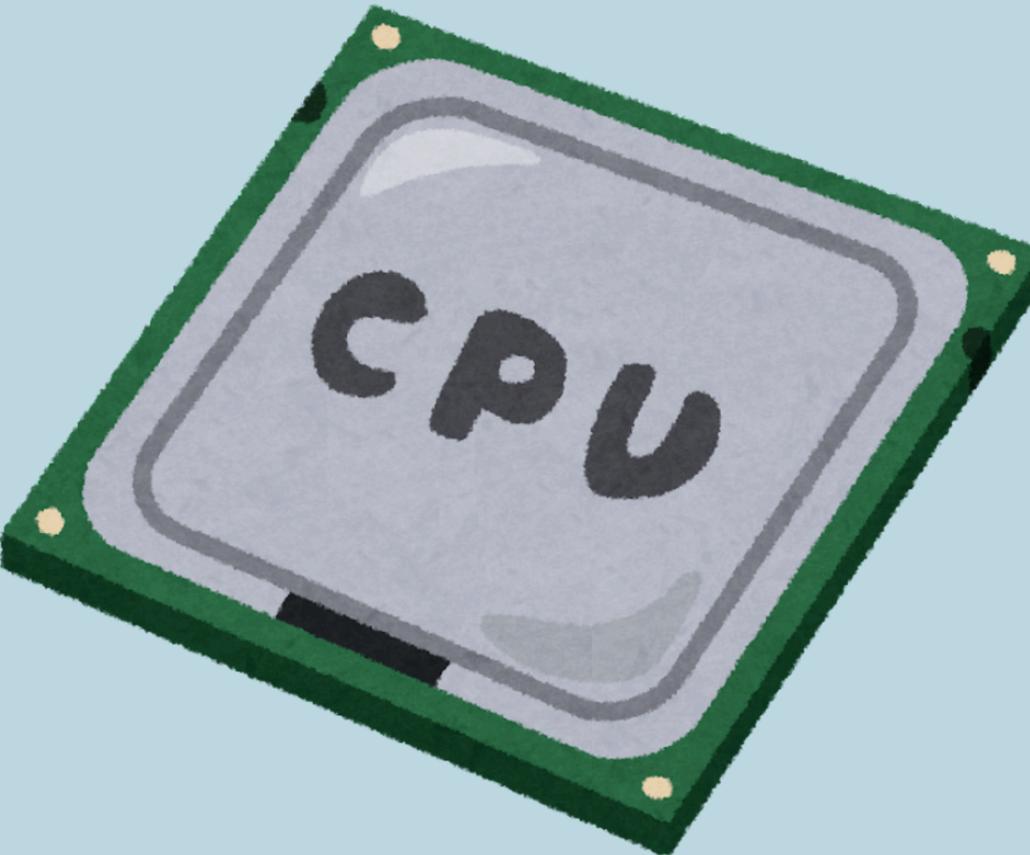
- Entrenamiento:

- El modelo se pone en entrenamiento con `model.train()`.
- Cada minibatch de imágenes y máscaras pasan por el modelo
- Se calcula la perdida y se hace back propagation con `los.backward()`
- El optimizador actualiza los pesos con `optimizar.step()`
- El Scheduler ajusta el learning rate con `scheduler.step()`

- Validación

- Después de cada epoch, el modelo se evalua en el dataset de validation con `model.eval()` sin actualizar pesos.
- Se calculan las métricas mIoU para ambos datasets.

CONSIDERACIONES DE HARDWARE



En kaggle hay distintas GPUs como:

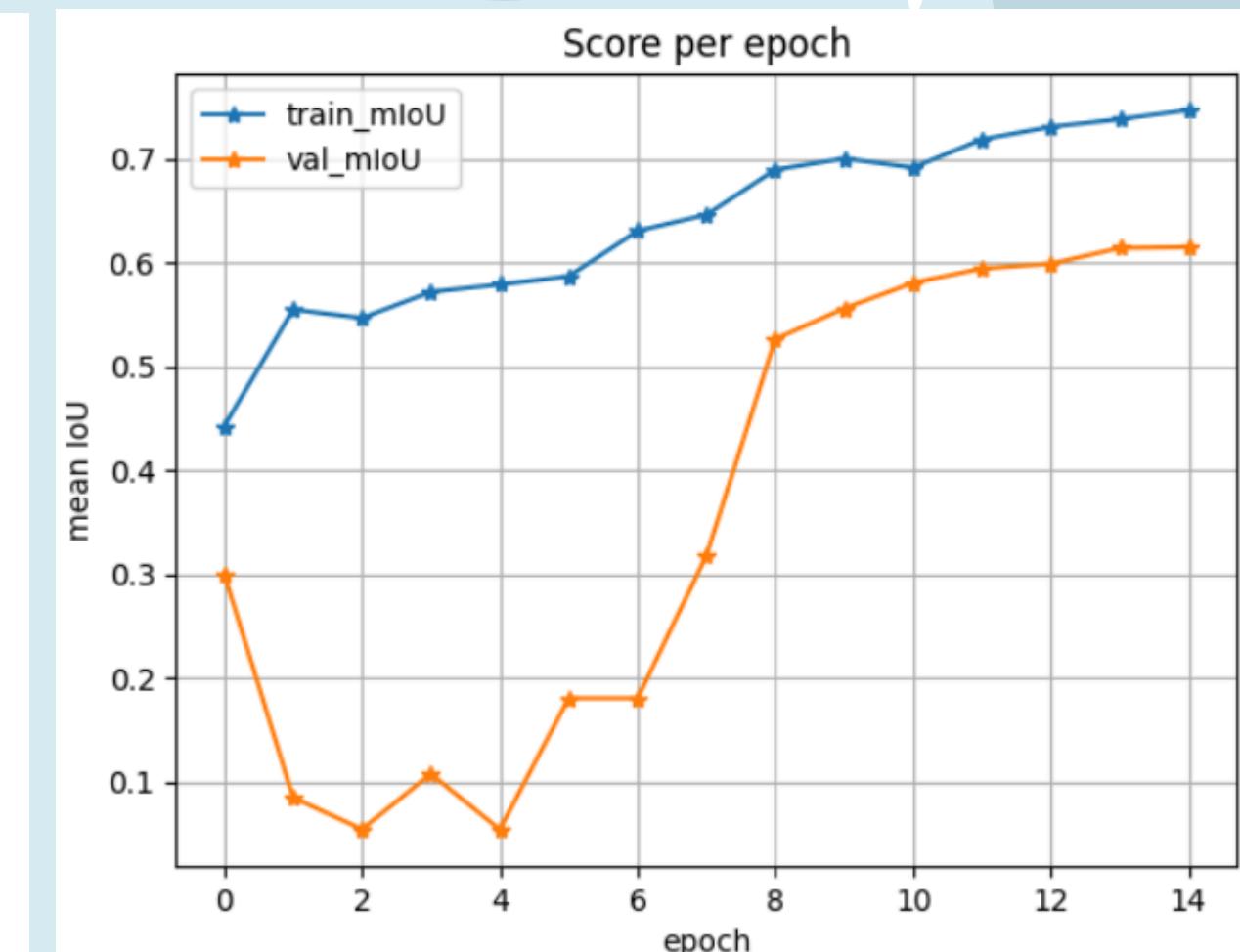
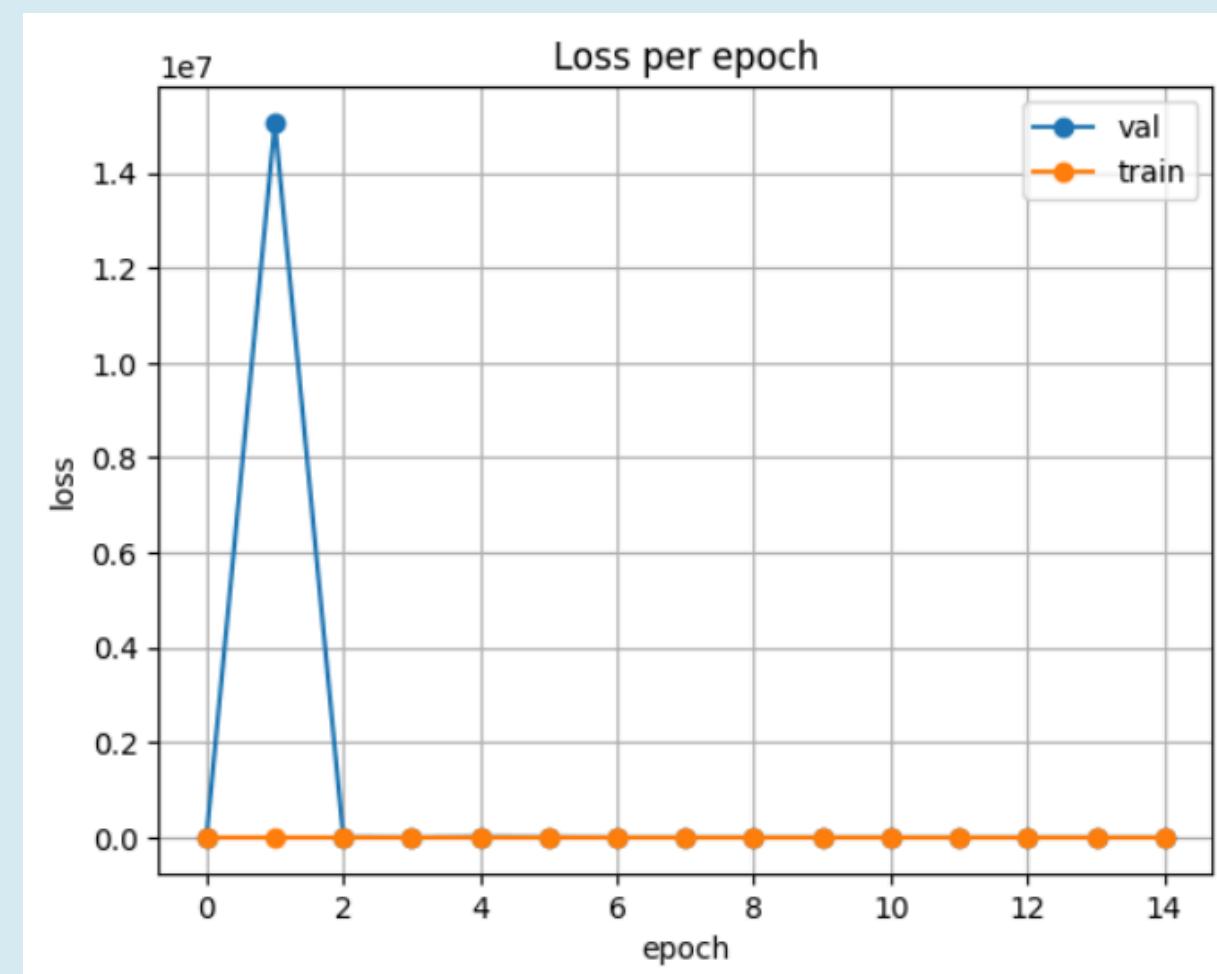
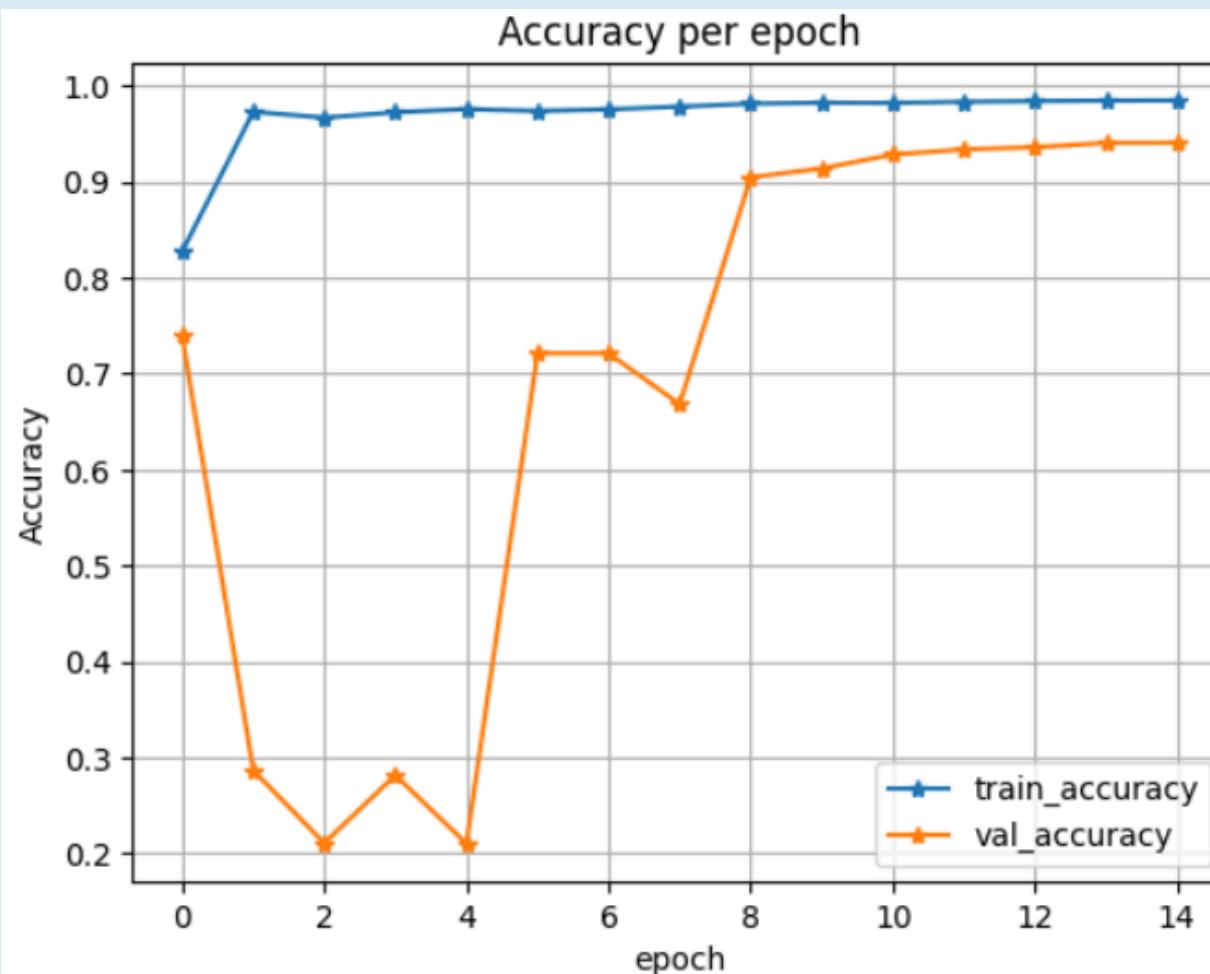
- T4
- P100
- V100

Con 16 GB de RAM son suficientes.

Nosotros usamos P100.

- Sin GPU, el tiempo de entrenamiento habría sido mucho mayor, dificultando la experimentación con múltiples configuraciones de hiperparámetros.

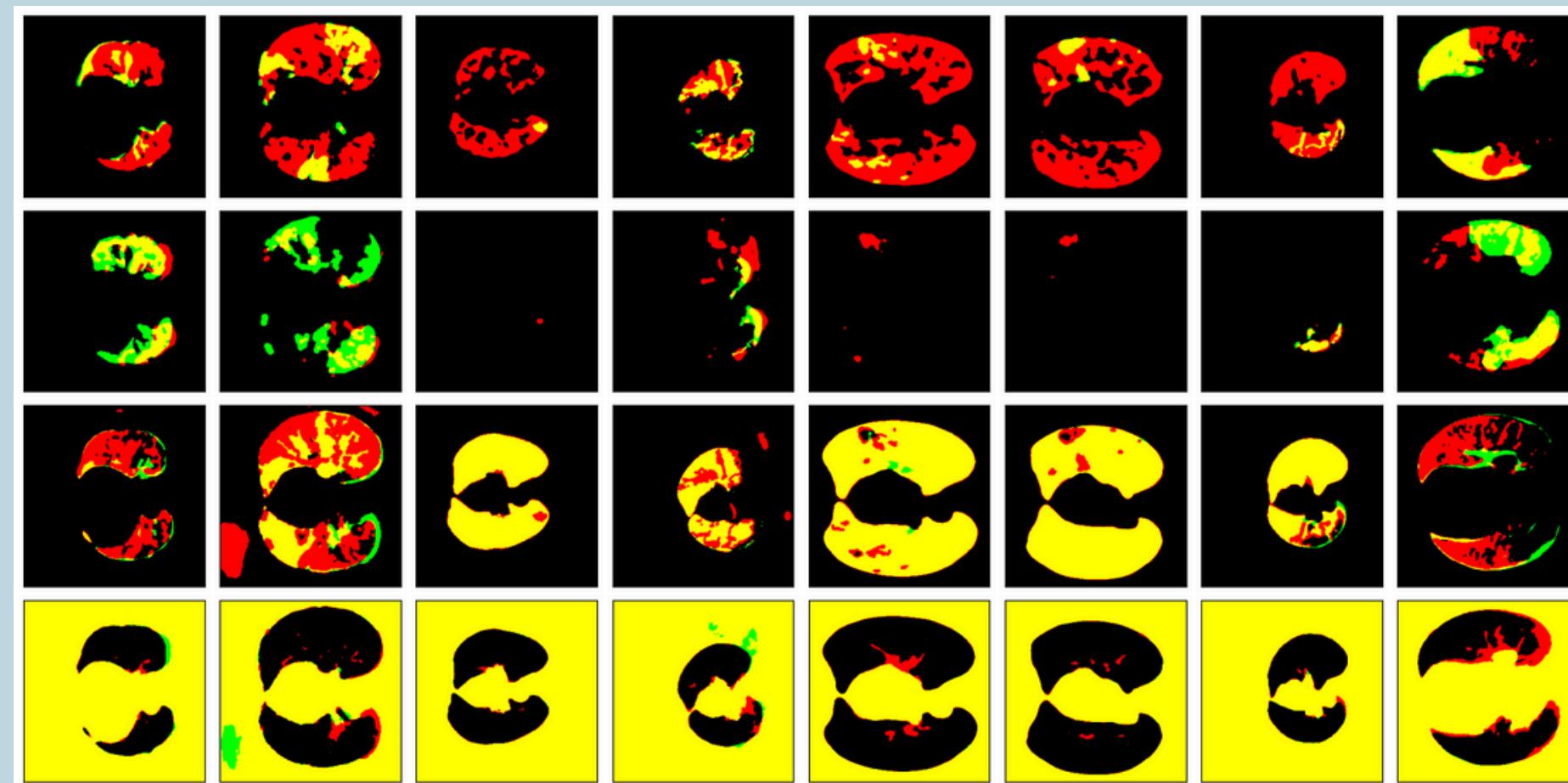
RESULTADOS DEL ENTRENAMIENTO



RESULTADOS



RESULTADOS - VISUALES



- Amarillo: Píxeles en donde tanto la predicción concuerda con la clase verdadera.
- Rojo: Píxeles en donde la predicción es positiva pero la máscara verdadera es negativa.
- Verde: Píxeles en donde la predicción es negativa pero la máscara verdadera es positiva.

💡 Podemos observar que hay una cantidad considerable de valores correctos con respecto a la verdad fundamental.

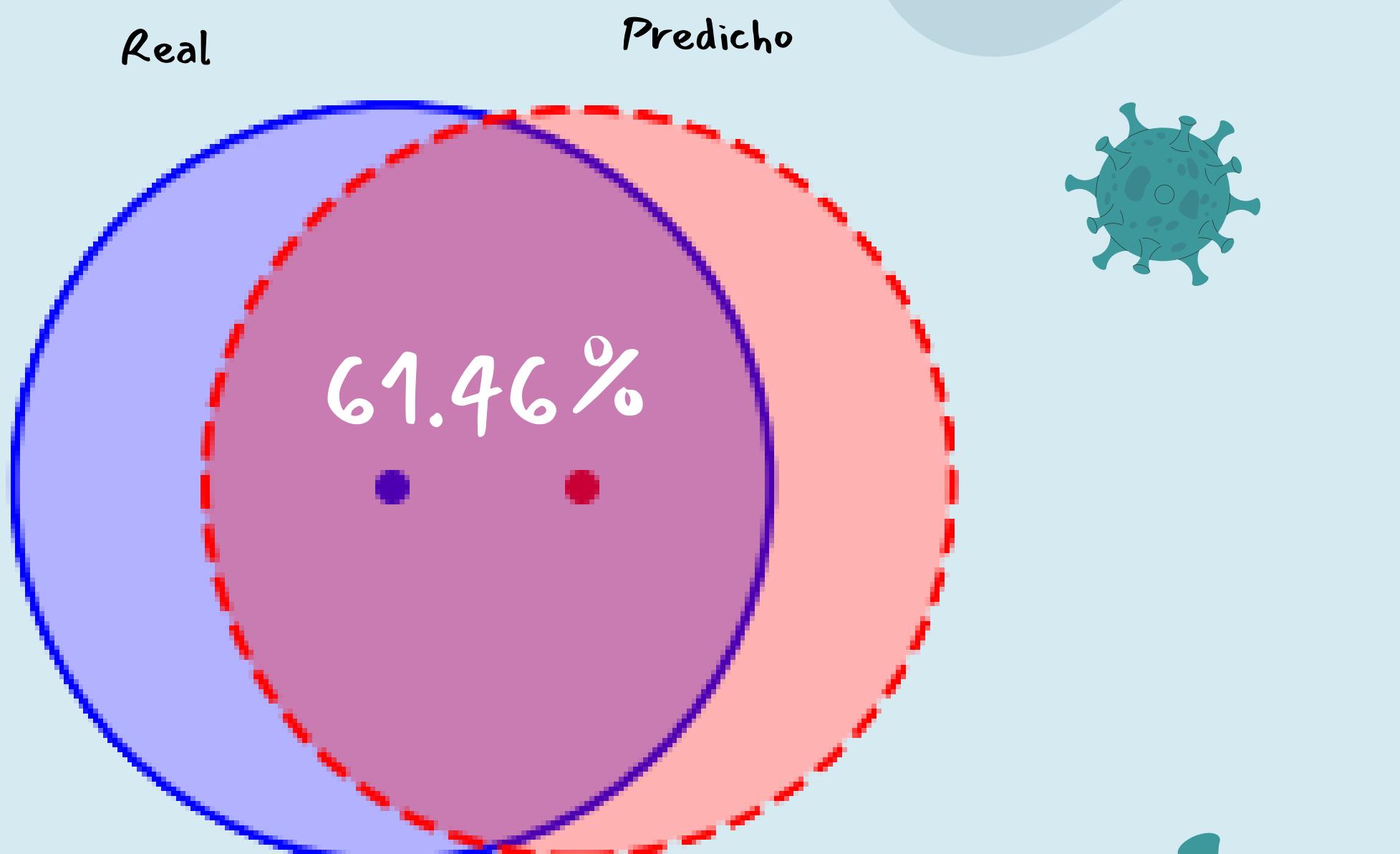
Sin embargo, hay varios píxeles rojos, lo que significa que se están identificando valores en una clase cuando en realidad no lo son.

RESULTADOS - CUANTITATIVOS

mIoU Score: 0.6146

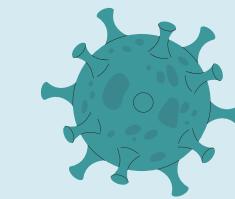
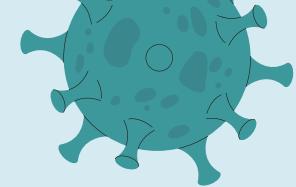
El mIoU de 0.6146 indica que el modelo delimita correctamente el 61% del área de los objetos.

Esto indica un desempeño suficiente para validar que la arquitectura puede detectar estas características críticas del COVID-19, pero presenta un margen considerable de mejora antes de ser confiable para asistir en diagnósticos médicos reales



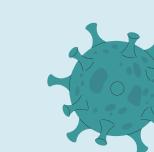
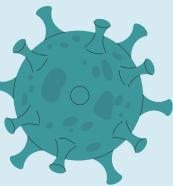
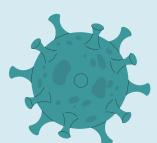
CONCLUSIÓN





REFERENCIAS

- Società Italiana di Radiologia Medica e Interventistica. (s.f.). Home. SIRM. Recuperado 14 de septiembre de 2024, de [Presentations are communication tools that can be used as demonstrations, lectures, speeches, reports, and more. Mostly presented before an audience, it serves a variety of purposes, making presentations powerful tools for convincing and teaching. To create a stunning presentation, it's best to simplify your thoughts.](#)
- Gaillard, F., & Weerakkody, Y. (s.f.). COVID-19. Radiopaedia.org. Recuperado 14 de septiembre de 2024, de [Presentations are communication tools that can be used as demonstrations, lectures, speeches, reports, and more. Mostly presented before an audience, it serves a variety of purposes, making presentations powerful tools for convincing and teaching. To create a stunning presentation, it's best to simplify your thoughts.](#)
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv. [Presentations are communication tools that can be used as demonstrations, lectures, speeches, reports, and more. Mostly presented before an audience, it serves a variety of purposes, making presentations powerful tools for convincing and teaching. To create a stunning presentation, it's best to simplify your thoughts.](#)
- GeekforGeeks. (s.f.). U-Net Architecture Explained. Recuperado 14 de septiembre de 2024, de [Presentations are communication tools that can be used as demonstrations, lectures, speeches, reports, and more. Mostly presented before an audience, it serves a variety of purposes, making presentations powerful tools for convincing and teaching. To create a stunning presentation, it's best to simplify your thoughts.](#)
- Pepe Cantoral, Ph.D. (2021, 25 agosto). Adam optimizador – Algoritmos de optimización para redes neuronales [Vídeo]. YouTube. <https://www.youtube.com/watch?v=QFys1W2NAAc>
- Pepe Cantoral, Ph.D. (2022, 11 julio). U-NET para segmentación semántica, explicación del paper. [Vídeo]. YouTube. <https://www.youtube.com/watch?v=waIPUsecaaQ>





“CUANTA MÁS LARGA
LA EXPLICACIÓN,
MÁS GRANDE LA
MENTIRA”

-- PROVERBIO CHINO

