

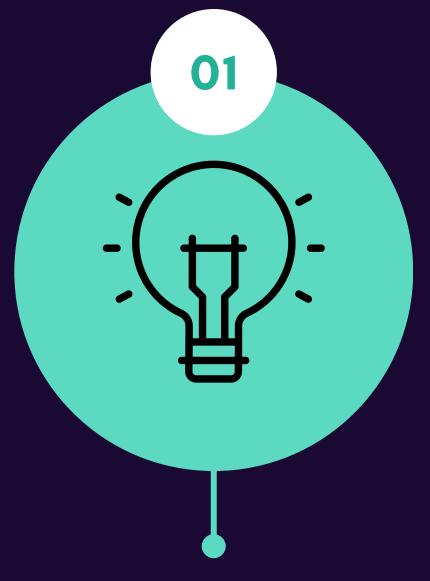
APRENDIZAJE PROFUNDO 1 Y 2

TUMORSYNTH: SÍNTESIS DE IMÁGENES DE TUMORES CEREBRALES CON GAN

Borja Esteve Molner

Máster en Inteligencia Artificial Avanzada y
Aplicada

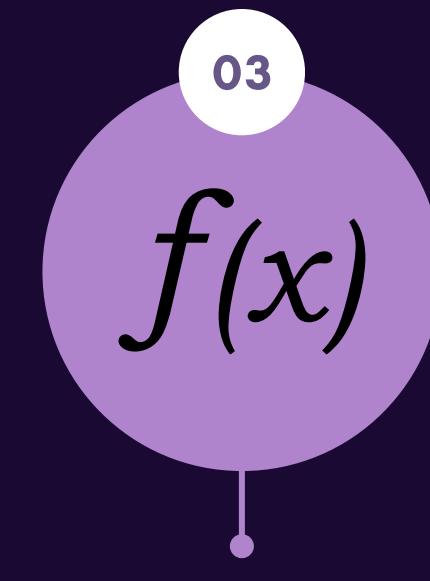
ÍNDICE



INTRODUCCIÓN



DATOS



MODELO



RESULTADO



CONCLUSIONES Y
TRABAJO FUTURO

INTRODUCCIÓN



TUMORES CEREBRALES

Los tumores cerebrales son crecimientos anormales de células en el cerebro, que pueden ser benignos o malignos. Pueden causar daño cerebral y ser potencialmente mortales. Se clasifican en primarios (originados en el cerebro) y secundarios (provenientes de otros órganos).



PROBLEMA

En España se detectan 5000 casos de tumores cerebrales al año. Esto hace muy complicado la aplicación de métodos de IA para la detección de estos tumores.



SOLUCIÓN

Los modelos generativos, son una clase de modelos de aprendizaje profundo que aprenden la distribución de datos subyacente de la muestra. Estos modelos se pueden utilizar para reducir datos a sus propiedades fundamentales o para generar nuevas muestras de datos con propiedades nuevas y variadas.

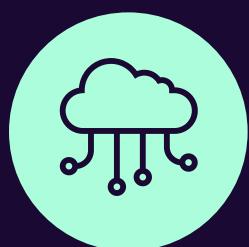
CONJUNTO DE DATOS

Las características del conjunto de datos son las siguientes:



RESONANCIAS MAGNÉTICAS

El dataset está compuesto por resonancias magnéticas del cerebro divididas en dos clases: YES y NO



CARACTERÍSTICAS

Compuesto por 253 imágenes:

- 155 de la clase YES
- 98 de la clase NO

Tamaños no regulares

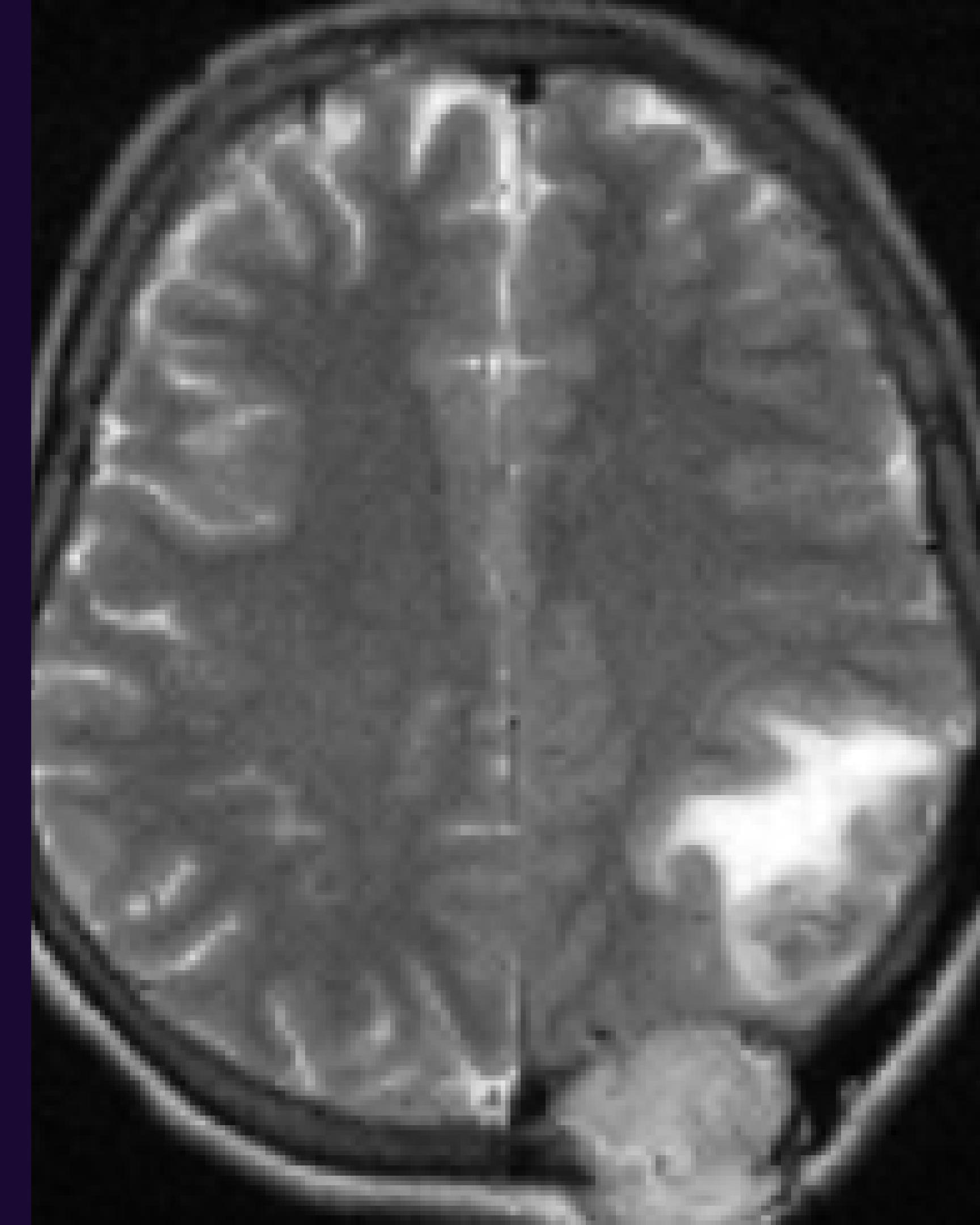
Un solo canal (imágenes en escala de grises)



PREPROCESADO

Red 1 y 2:

- Redimensión a 128x128
- Normalización de las imágenes



MODELOS

Los dos modelos se fundamentan en arquitecturas GAN.

01

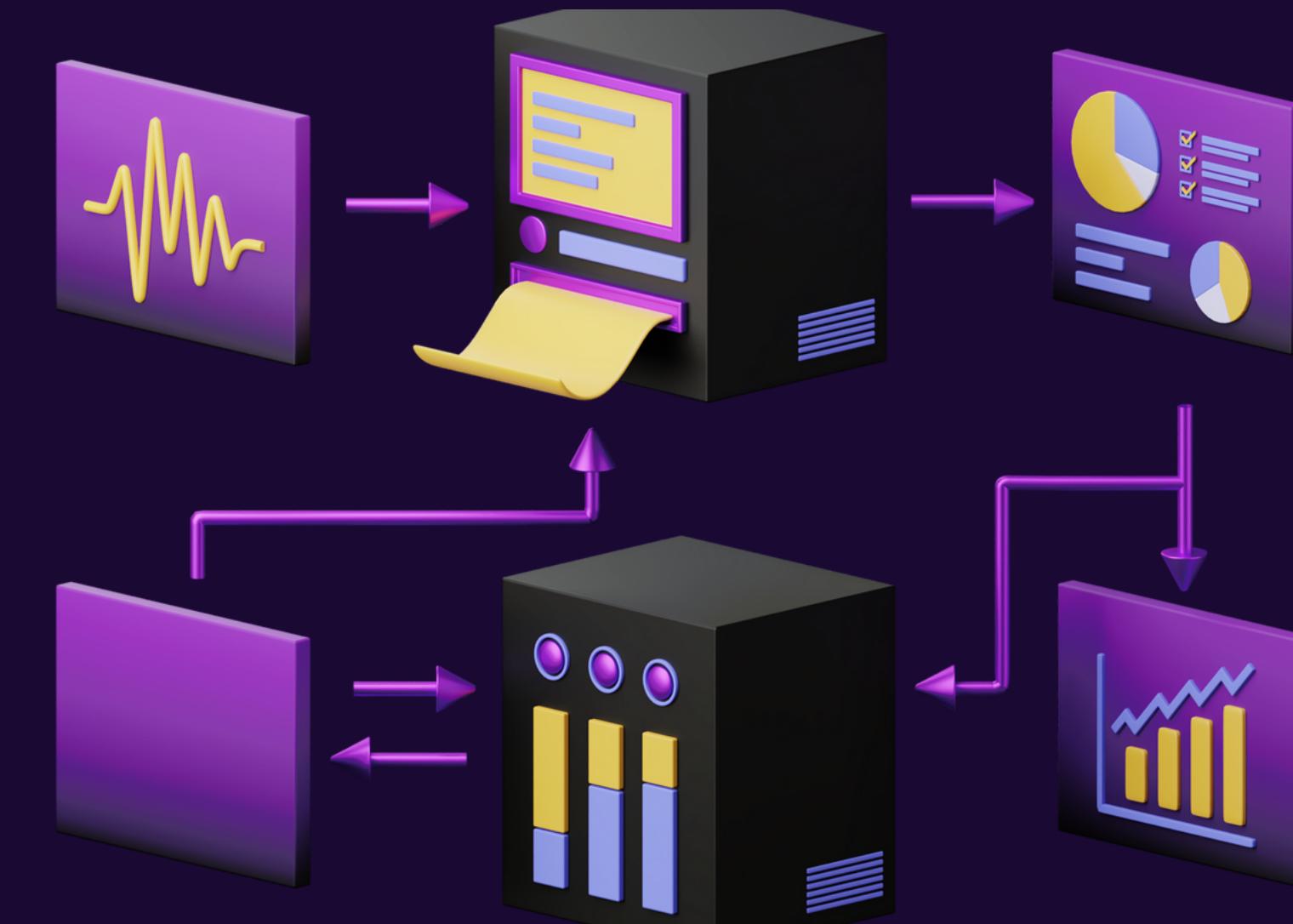
TUMORSYNTH 1

Modelo con el generador más simple.
Mismo discriminador que TUMORSYNT2.

02

TUMORSYNTH 2

Modelo con un generador con
arquitectura encoder-decoder más
sotisficado que TUMORSYNT2.



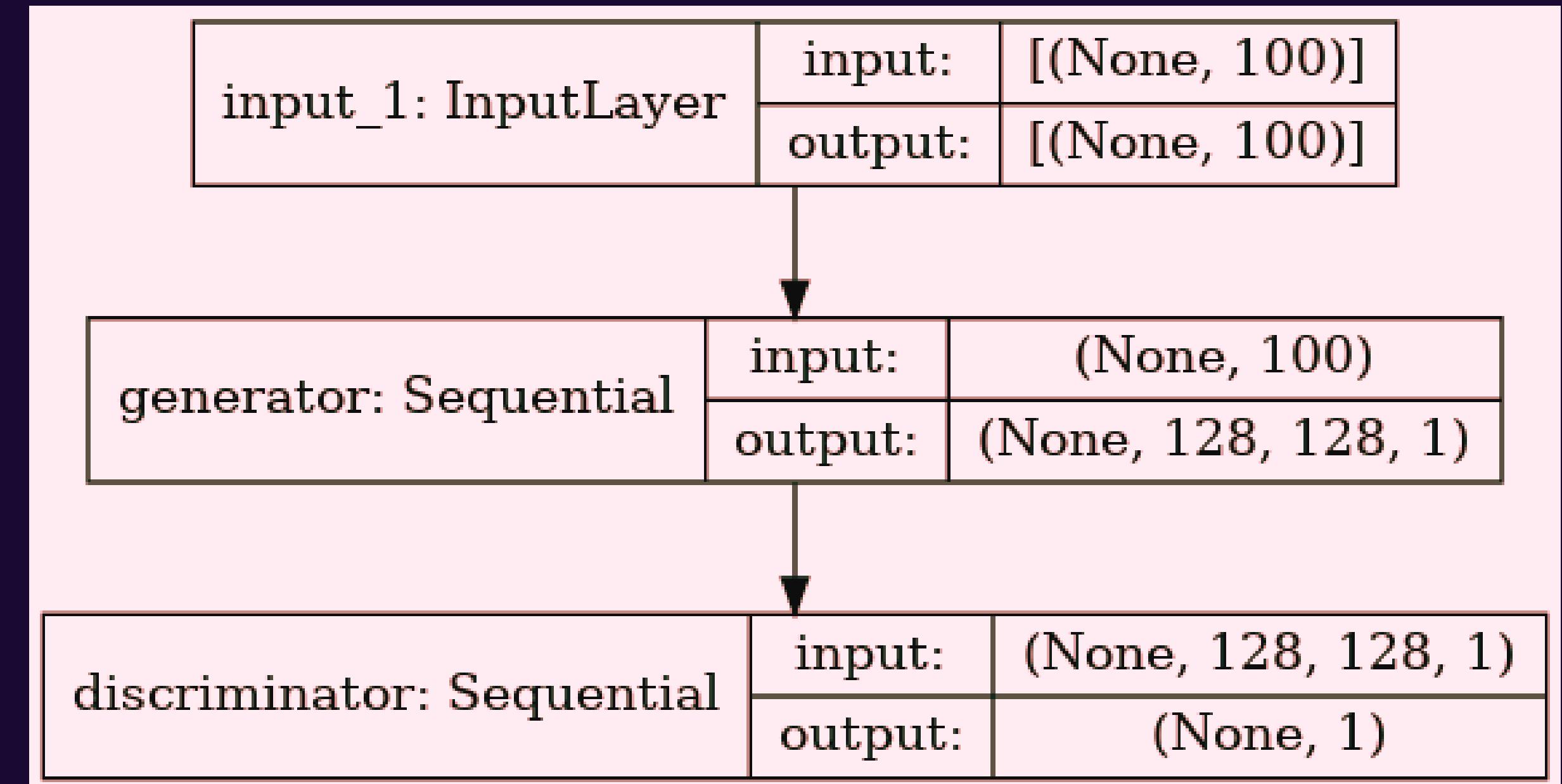
TUMORSYNTH1

27,848,066

Parámetros entrenables

PARÁMETROS

- Tamaño del ruido: 100
- Tamaño del lote: 4
- Optimizador: Adam con una tasa de aprendizaje de 0.0002 y un valor beta_1 de 0.5.
- Función de pérdida: Entropía cruzada binaria (tanto en generador como en discriminador).
- Número de Epochs: 18
- Steps por Epoch: 3750
- Mejor época: 8



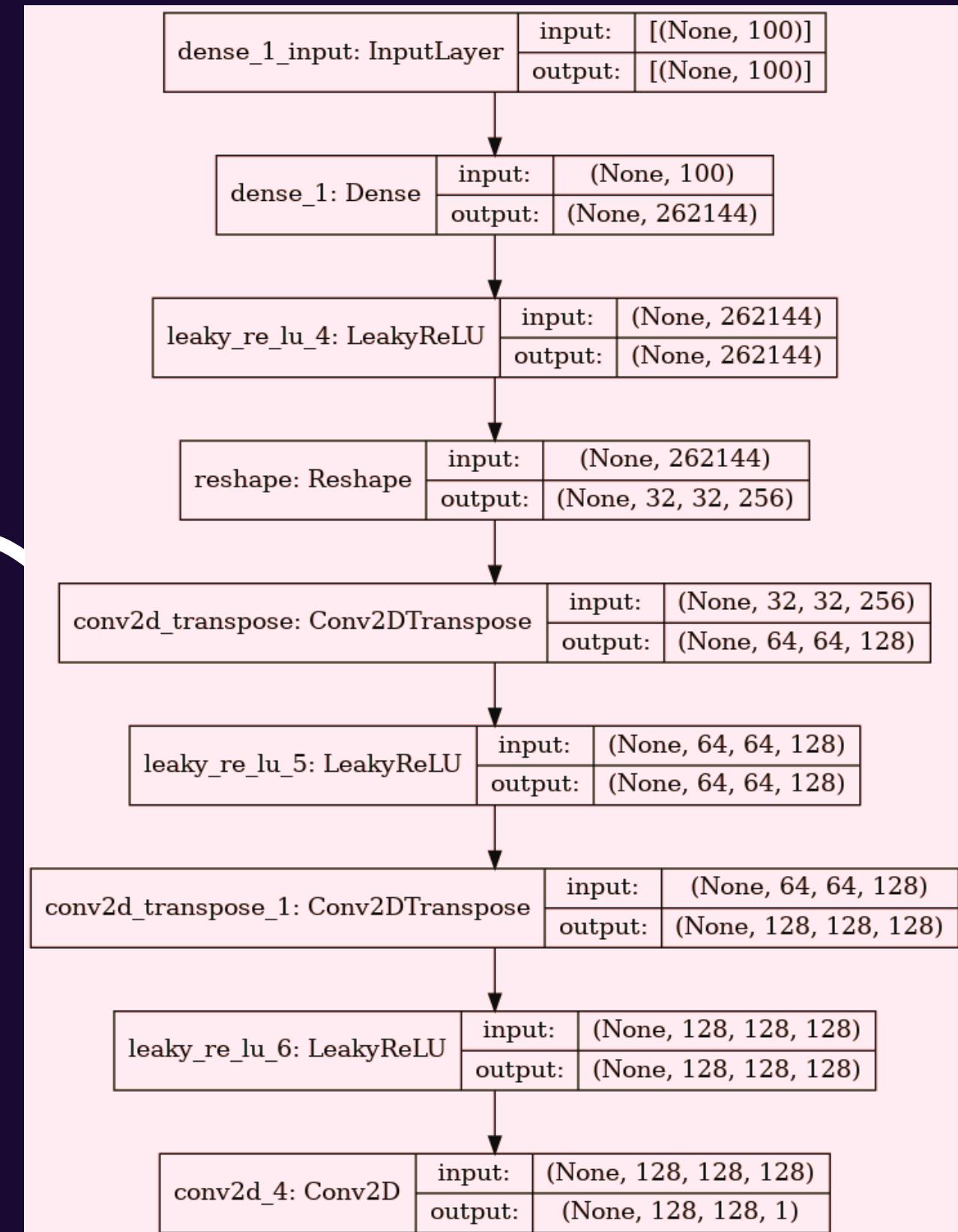
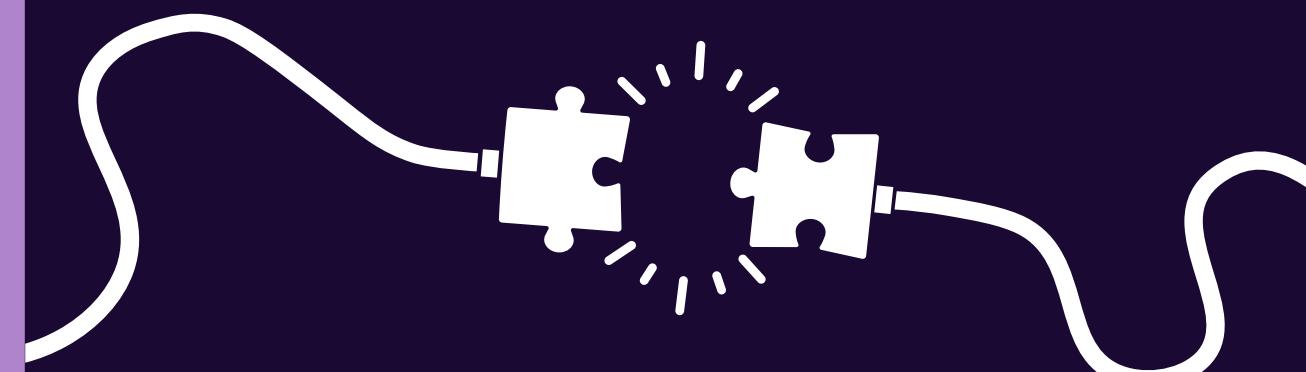
TUMORSYNTH1

27,848,066

Parámetros entrenables

GENERADOR

- Capa 1: Capa densa con 32x32x256 neuronas, seguida de activación LeakyReLU.
- Capa 2: Capa de red convolucional transpuesta con 128 filtros, kernel size (4,4), strides (2,2) y activación LeakyReLU.
- Capa 3: Capa de red convolucional transpuesta con 128 filtros, kernel size (4,4), strides (2,2) y activación LeakyReLU.
- Capa de salida: Capa de convolución con 1 filtro, kernel size (4,4), activación tangente hiperbólica (tanh)

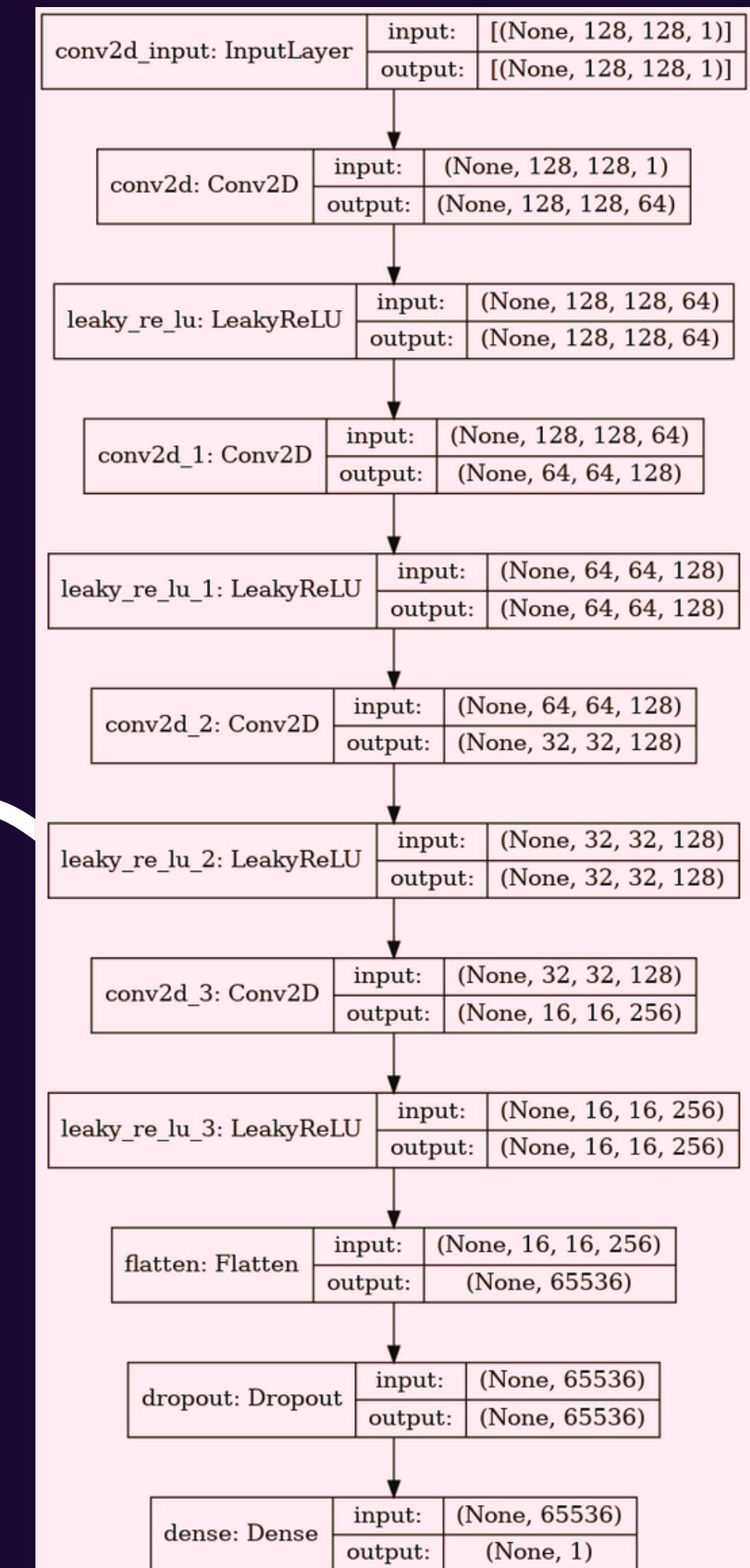
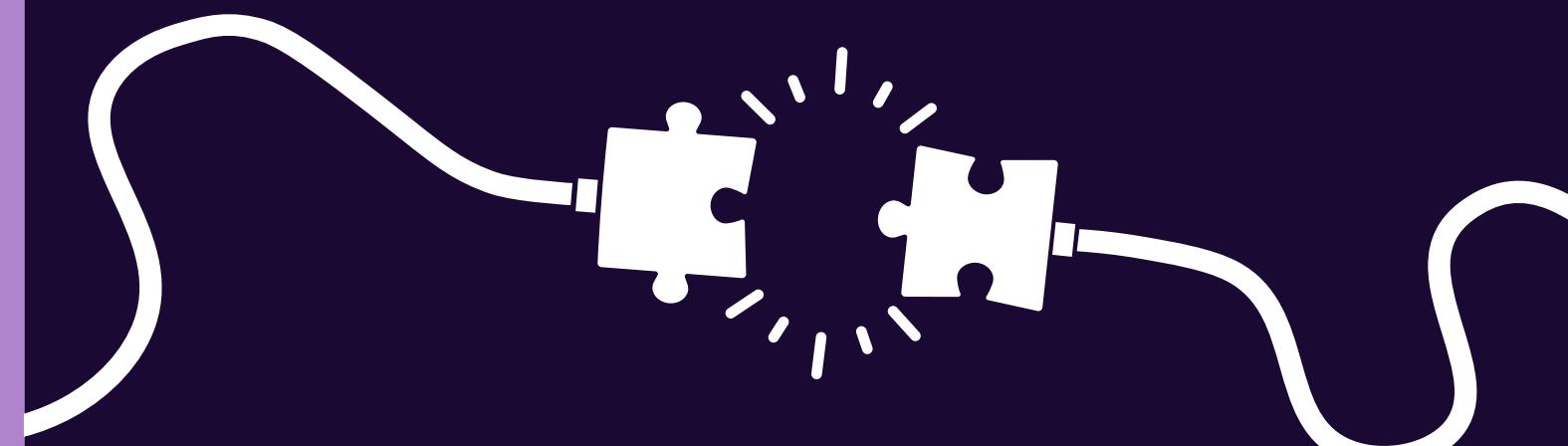


TUMORSYNTH1

27,848,066
Parámetros entrenables

DISCRIMINADOR

- Capa 1: Capa de convolución con 64 filtros, kernel size (3,3) y activación LeakyReLU.
- Capa 2: Capa de convolución con 128 filtros, kernel size (3,3), strides (2,2) y activación LeakyReLU.
- Capa 3: Capa de convolución con 128 filtros, kernel size (3,3), strides (2,2) y activación LeakyReLU.
- Capa 4: Capa de convolución con 256 filtros, kernel size (3,3), strides (2,2) y activación LeakyReLU.
- Capa de aplanamiento (Flatten).
- Capa de dropout con una tasa del 40%.
- Capa de salida densa con activación sigmoide para la clasificación binaria.



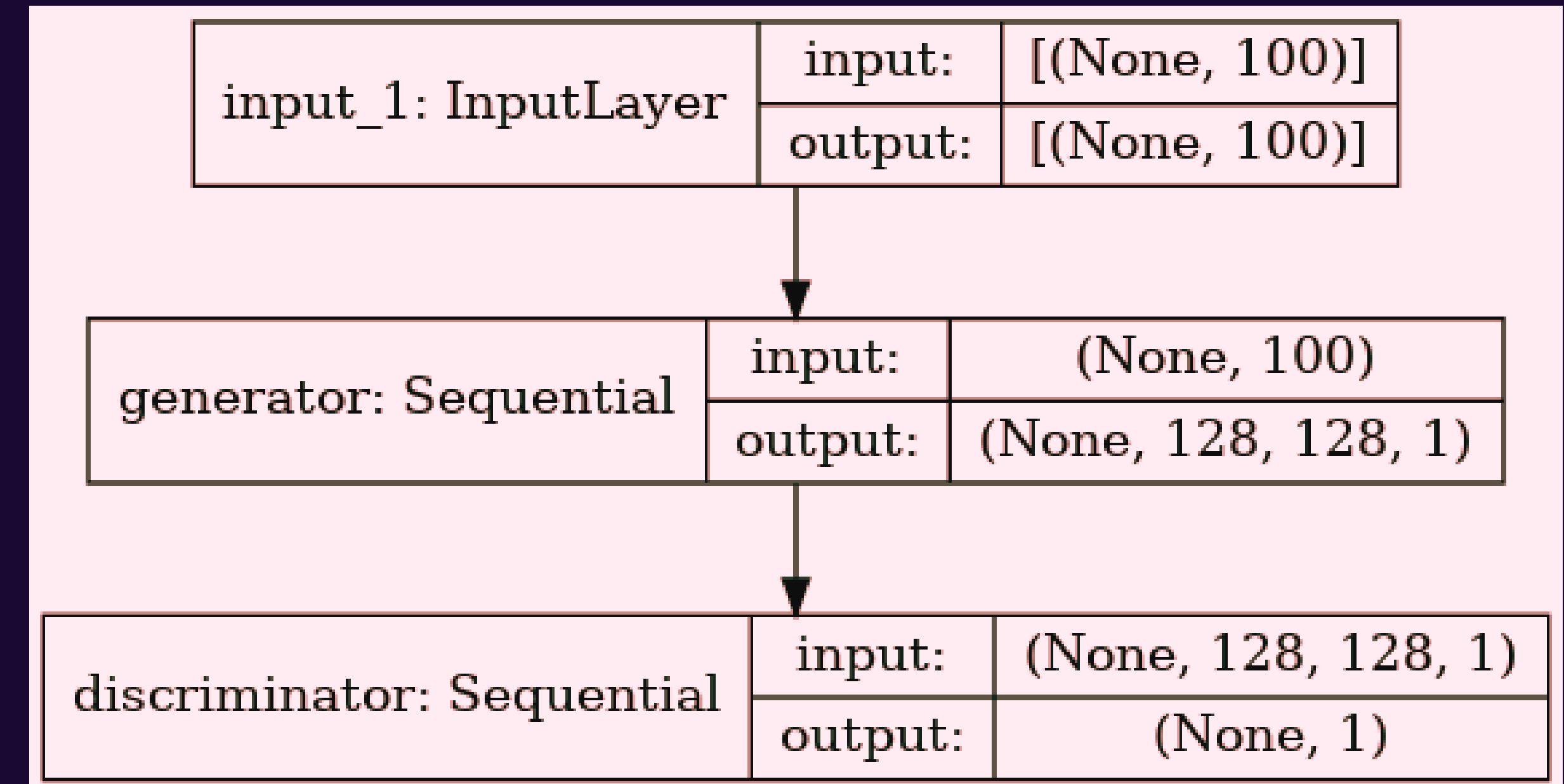
TUMORSYNTH2

427,362,946

Parámetros entrenables

PARÁMETROS

- Tamaño del ruido: 100
- Tamaño del lote: 4
- Optimizador: Adam con una tasa de aprendizaje de 0.0002 y un valor beta_1 de 0.5.
- Función de pérdida: Entropía cruzada binaria (tanto en generador como en discriminador).
- Número de Epochs: 30
- Steps por Epoch: 3750
- Mejor época: 27



TUMORSYNTH2

427,362,946

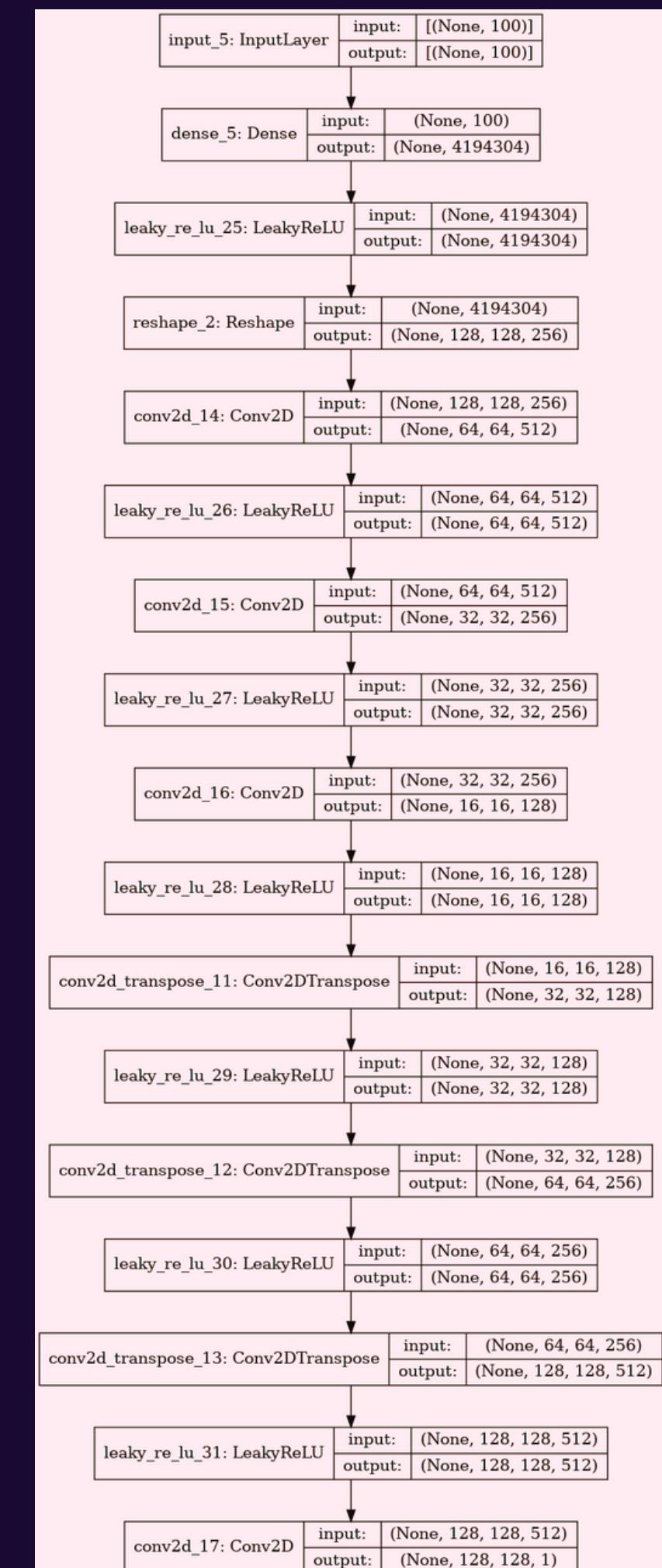
Parámetros entrenables

GENERADOR

- Capa densa con 128x128x256 neuronas, seguida de activación LeakyReLU.
- Encoder:
 - a. Tres capas convolucionales
 - b. Tamaños de filtro 512, 256 y 128
 - c. Kernel size de 3x3 y Stride de 2x2
 - d. Activaciones LeakyReLU después de cada capa

GENERADOR

- Decoder:
 - a. Tres capas de convolución Transpuesta
 - b. Tamaños de filtro de 128, 256 y 512
 - c. Tamaño de kernel de 2x2 y Stride de 2x2
 - d. Activaciones LeakyReLU después de cada capa
- Capa de salida: Capa de convolución con 1 filtro, kernel size (4,4), activación tangente hiperbólica (tanh)

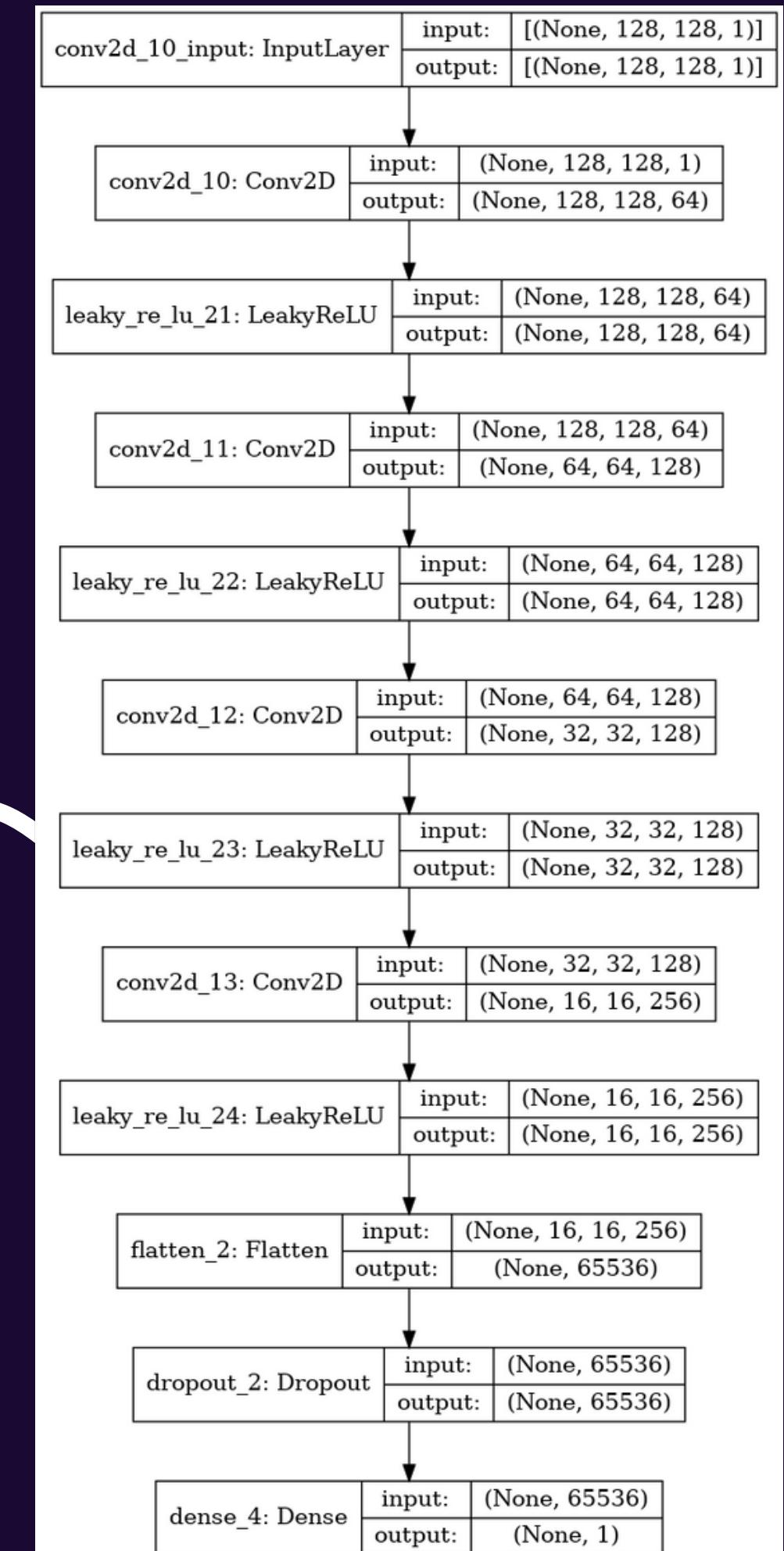
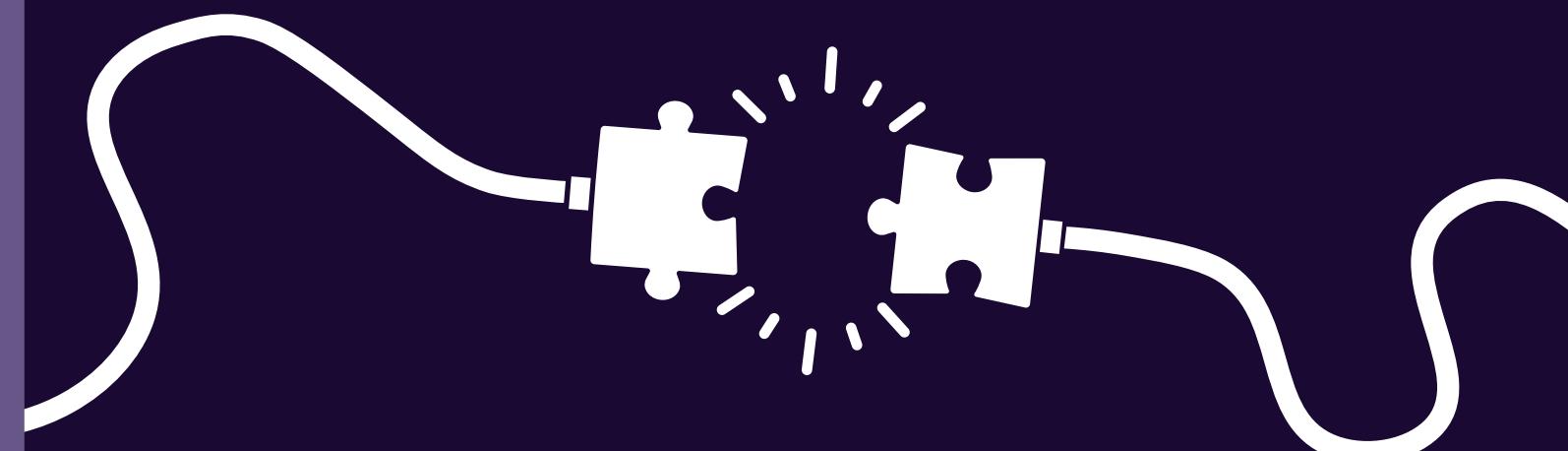


TUMORSYNTH2

427,362,946
Parámetros entrenables

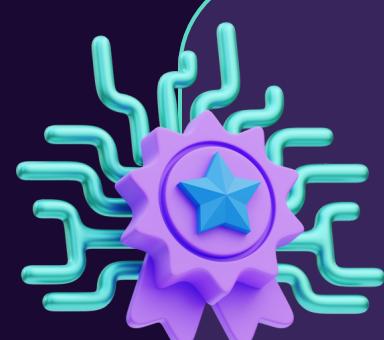
DISCRIMINADOR

- Capa 1: Capa de convolución con 64 filtros, kernel size (3,3) y activación LeakyReLU.
- Capa 2: Capa de convolución con 128 filtros, kernel size (3,3), strides (2,2) y activación LeakyReLU.
- Capa 3: Capa de convolución con 128 filtros, kernel size (3,3), strides (2,2) y activación LeakyReLU.
- Capa 4: Capa de convolución con 256 filtros, kernel size (3,3), strides (2,2) y activación LeakyReLU.
- Capa de aplanamiento (Flatten).
- Capa de dropout con una tasa del 40%.
- Capa de salida densa con activación sigmoide para la clasificación binaria.



MODELOS

Otros enfoques probados sin éxito



SAGAN - Self Attention GAN

Se infiere que la cantidad de imágenes en el conjunto no era suficiente para que la red tuviera un buen rendimiento.



VGG16 GAN

Se intentó emplear la arquitectura VGG16 como un extractor de características, pero los resultados no fueron satisfactorios.



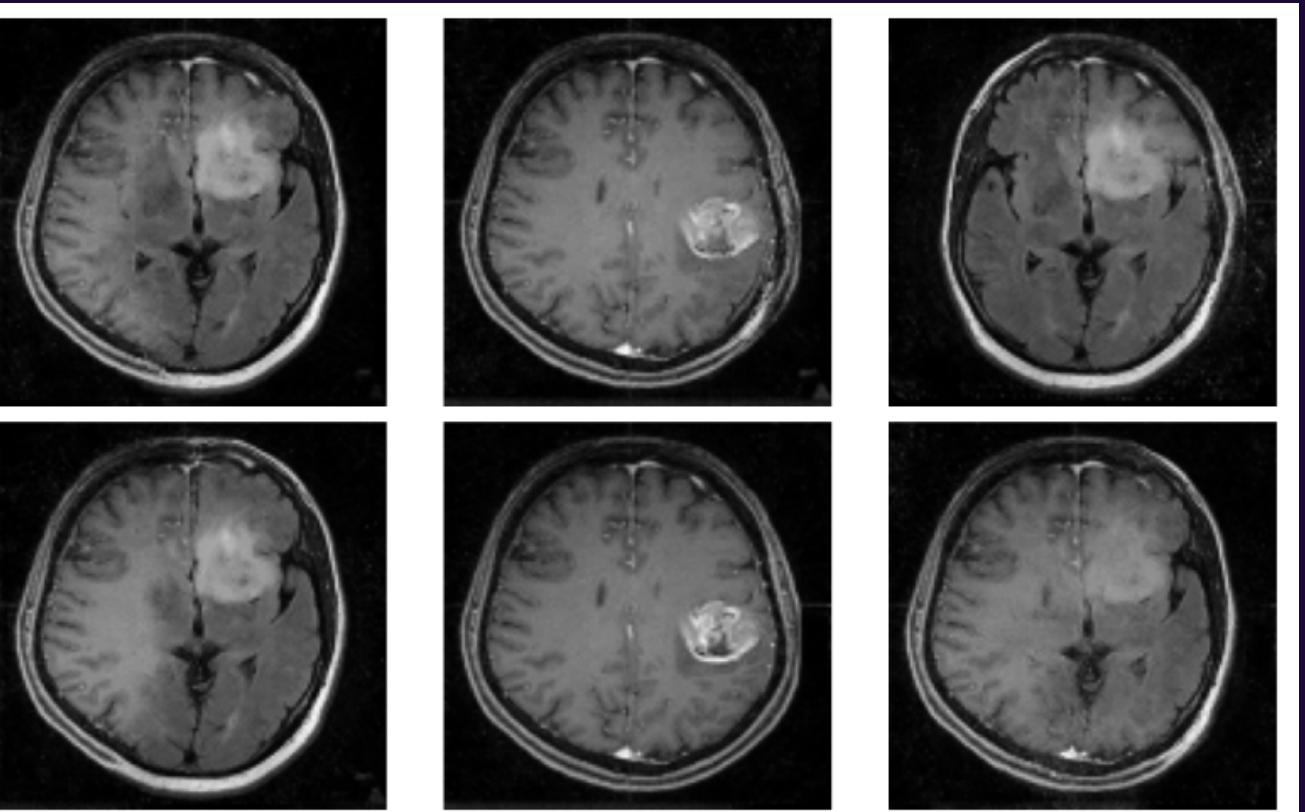
Aumento de datos

Se intentó aplicar un enfoque de preprocesamiento más complejo, pero no se obtuvieron los resultados deseados.

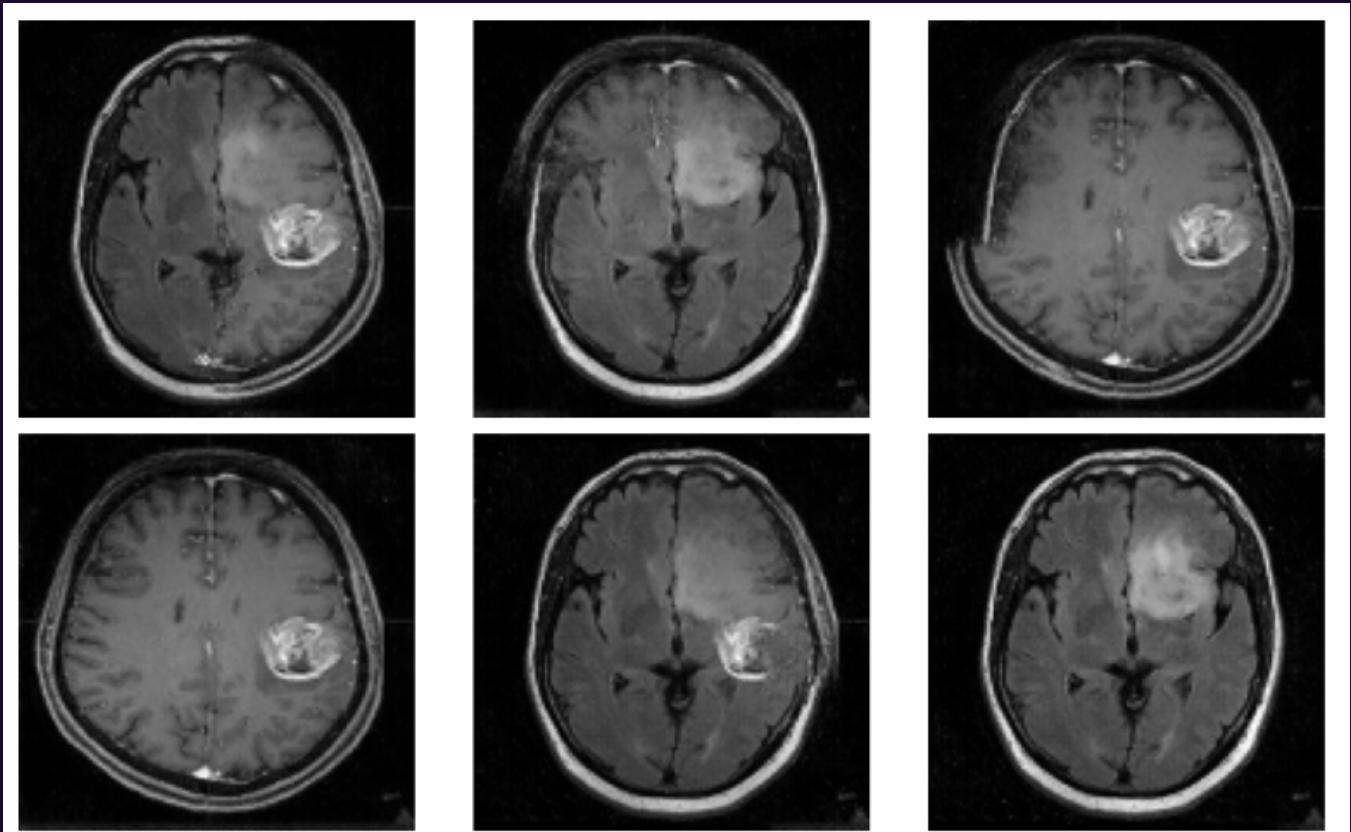
RESULTADOS

A continuación se muestran los resultados de ambas redes

TUMORSYNTH 1



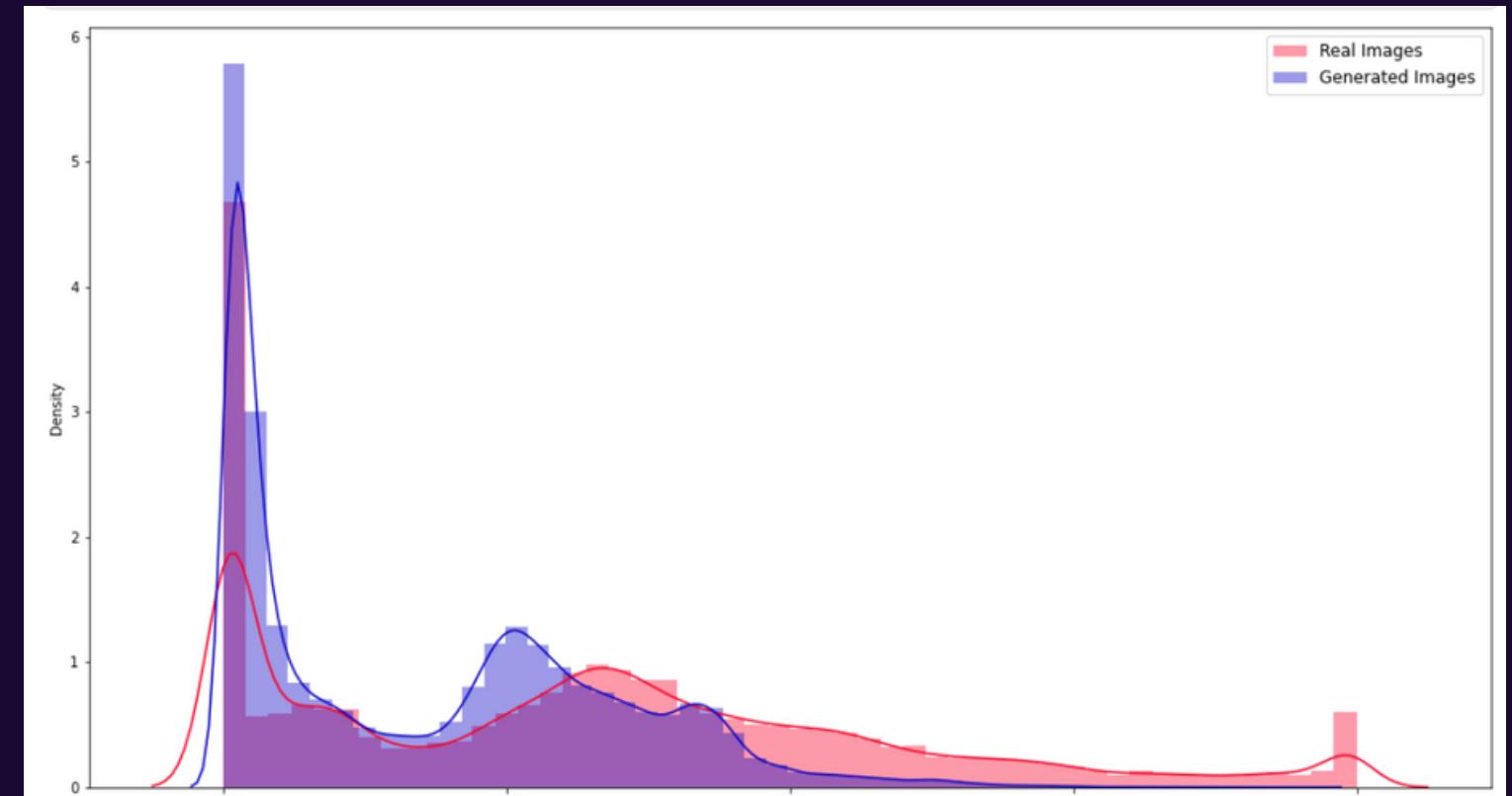
TUMORSYNTH 2



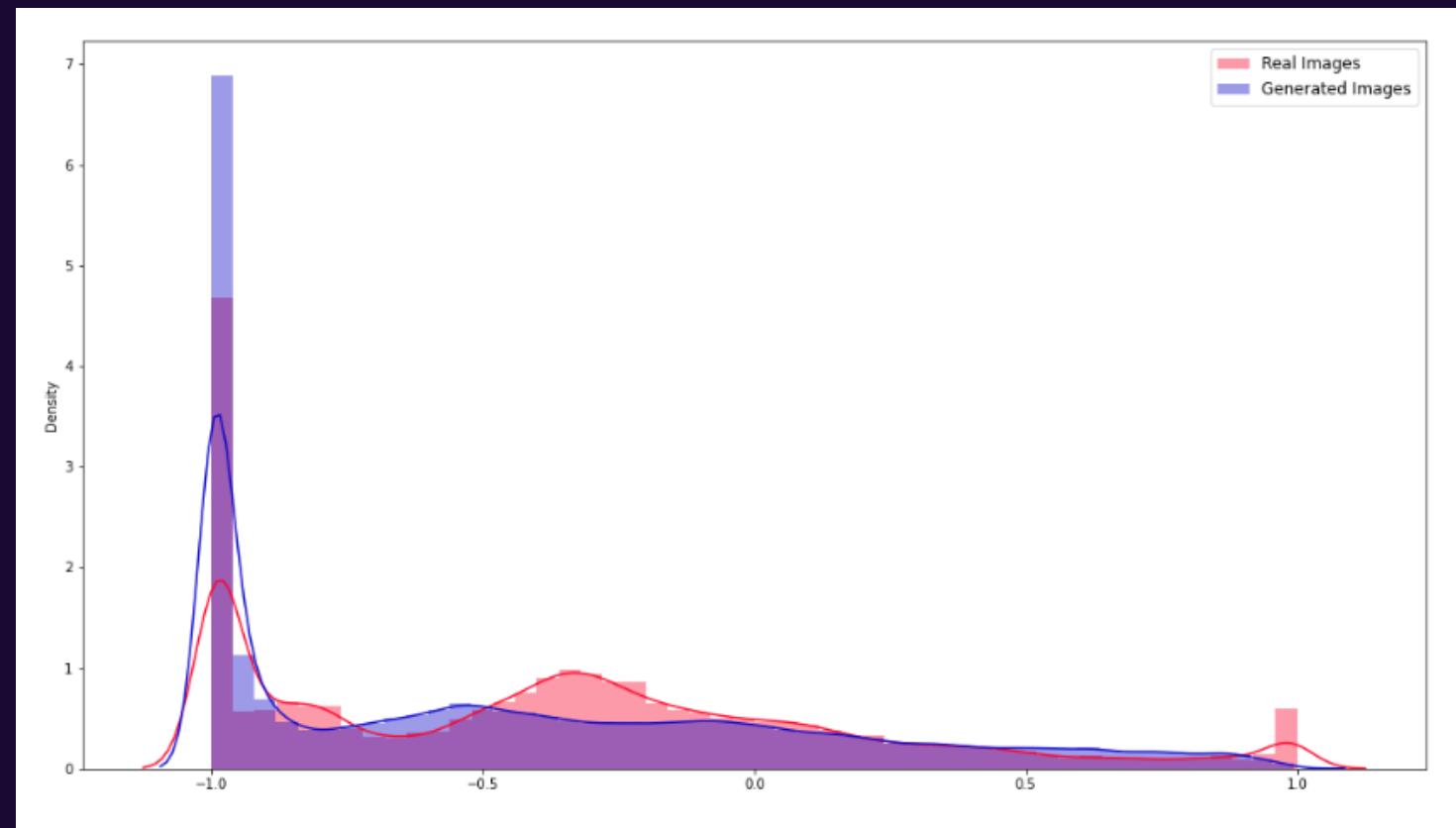
RESULTADOS

A continuación se muestran los resultados de ambas redes

TUMORSYNTH 1



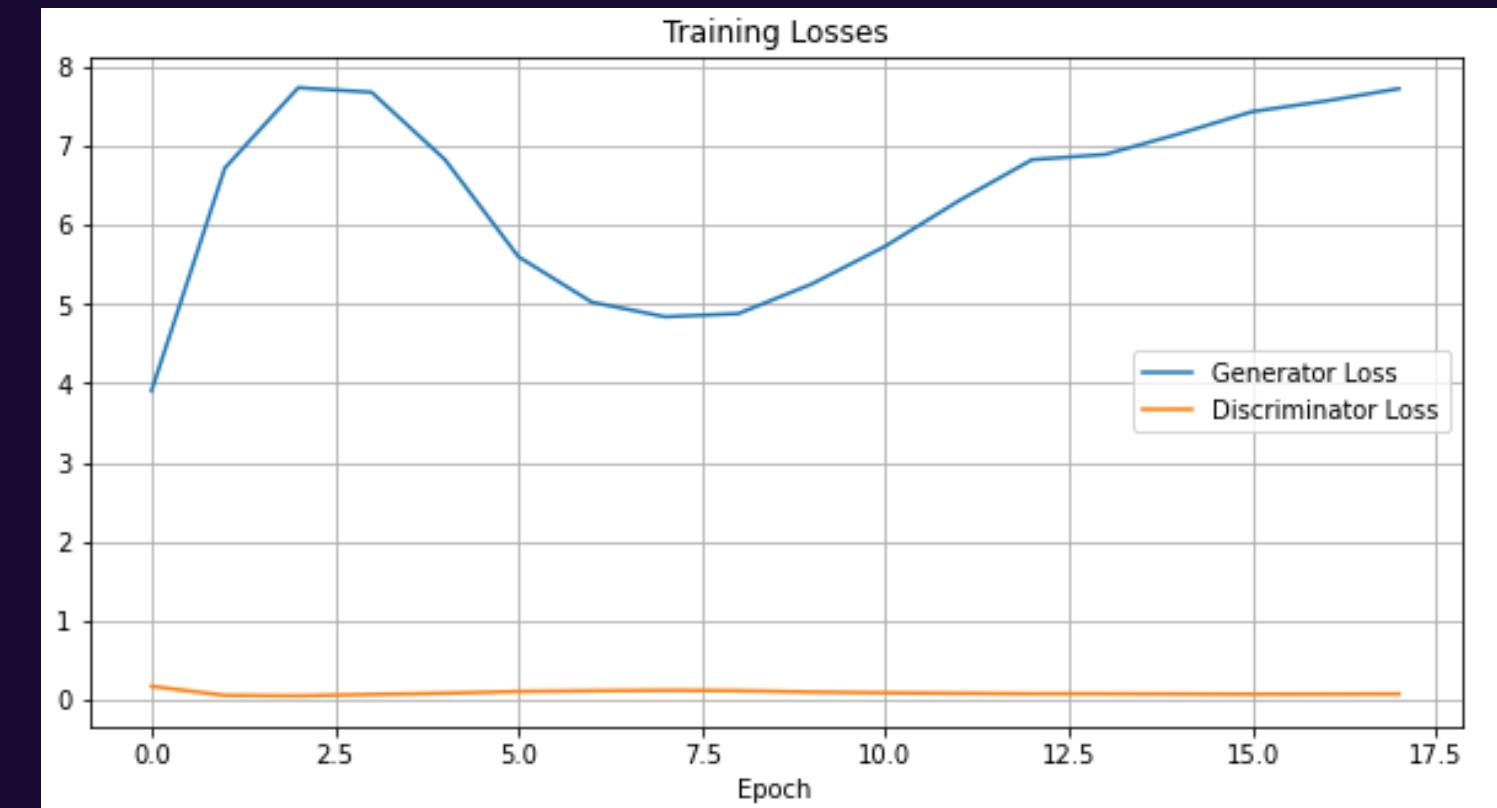
TUMORSYNTH 2



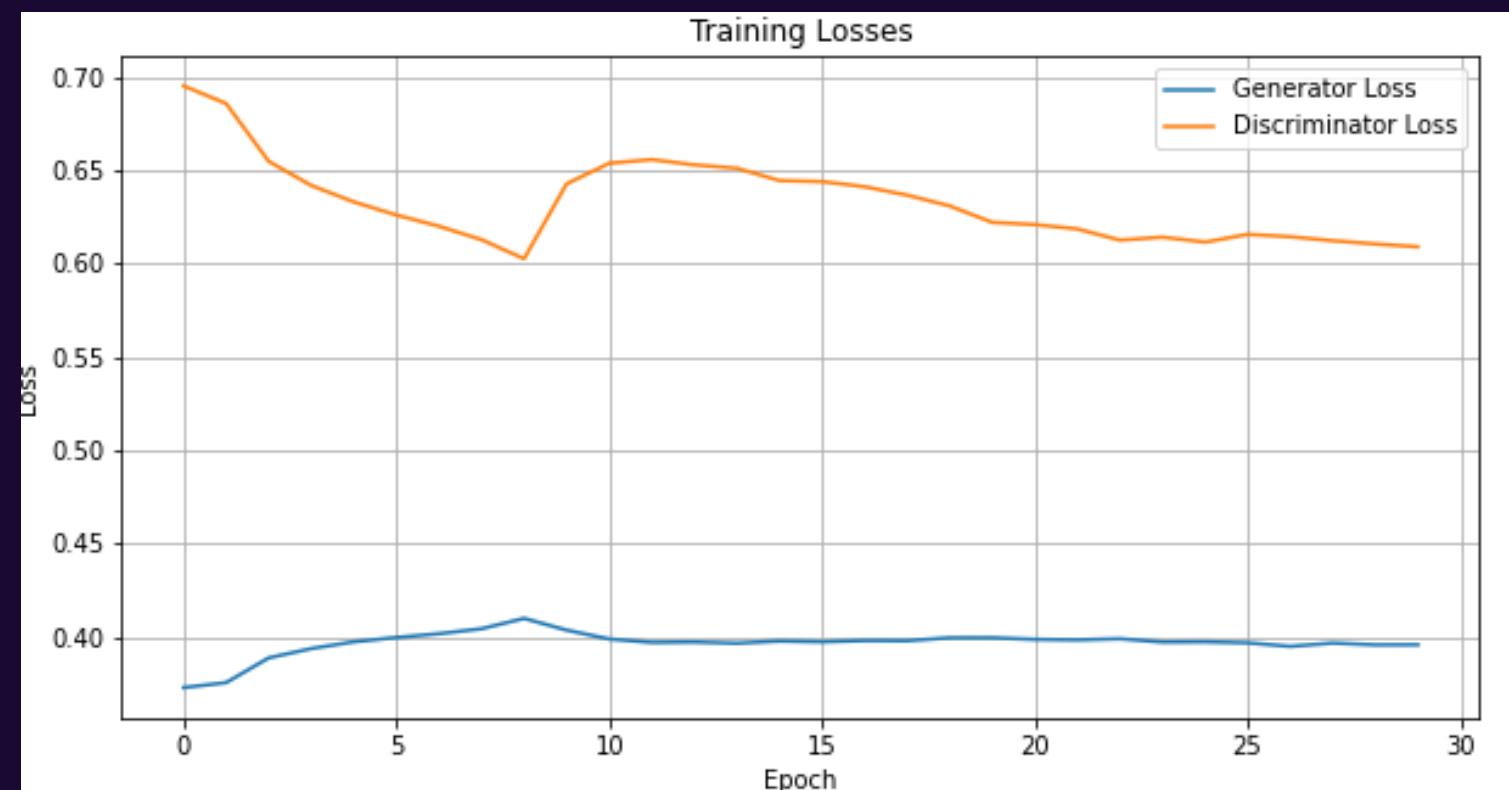
RESULTADOS

A continuación se muestran los resultados de ambas redes

TUMORSYNTH 1



TUMORSYNTH 2



CONCLUSIONES

Algunas de las conclusiones extraídas de la investigación son...

01

COMPLEJIDAD DEL MODELO

La red con un generador basado en un encoder-decoder captura con más fidelidad la distribución de probabilidad de las imágenes reales.

02

MODO COLAPSO

Aunque la red con el generador más simple parece generar imágenes más realistas a simple vista, es posible que esté experimentando un fenómeno conocido como "modo colapso"

03

DISTRIBUCIÓN DE PROBABILIDAD

Aunque la distribución de probabilidad de las imágenes generadas por el modelo complejo se ajusta mejor a las imágenes reales, esto no necesariamente se traduce en una mayor calidad visual de las muestras generadas.

04

ANÁLISIS VISUAL

A pesar de la discrepancia en la distribución de probabilidad, el análisis visual es una herramienta crucial para evaluar la calidad y larealismo de las imágenes generadas.

05

EQUILIBRIO ENTRE COMPLEJIDAD Y CAPACIDAD DE GENERALIZACIÓN

El modelo más complejo puede estar sobreajustado, por ello se ajusta bien a la distribución de probabilidad pero visualmente es peor.

TRABAJO FUTURO

Algunas de las conclusiones extraídas de la investigación son...

01

UTILIZACIÓN DE MÉTRICAS DE EVALUACIÓN ALTERNATIVAS

Evaluar la red con otras métricas como el Inception Score o el FID para mayor información.

02

ARQUITECTURAS

Probar diferentes arquitecturas como generador basado en ResNet, Unet...

03

DATOS

Utilización de una base de datos con mayor número de imágenes.



**¡GRACIAS POR
LA ATENCIÓN!**

BORJA ESTEVE MOLNER