



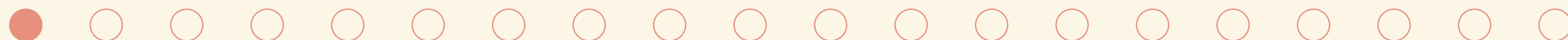
UNIVERSIDAD PERUANA
CAYETANO HEREDIA

Optimized U-Net for Brain Tumor Segmentation

Aplicaciones Clínicas en Señales e Imágenes

Grupo 1

Semestre 2024-2



Contenido

Avance 1 | Entendimiento de la data

- 01** Introducción y Problemática
- 02** Base de Datos
- 03** Metodología
- 04** Resultados



Introducción

La segmentación tumoral...

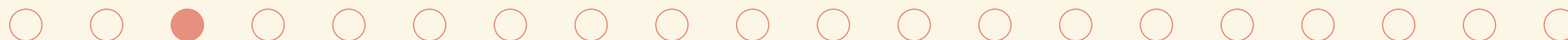


- Identificar
- Delinear regiones afectadas por un tumor

La segmentación automática



- ✓ Proporciona asistencia a los radiólogos
- ✓ Enfoque más preciso y fiable
- ✓ Permite planificación de tratamiento y seguimiento con más tiempo



Problemática

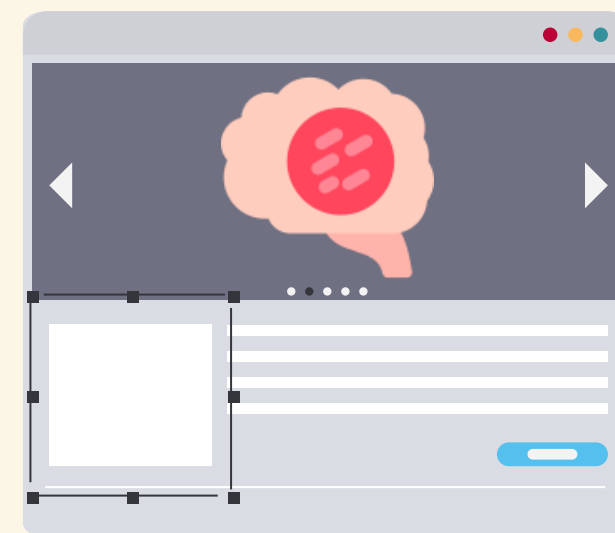
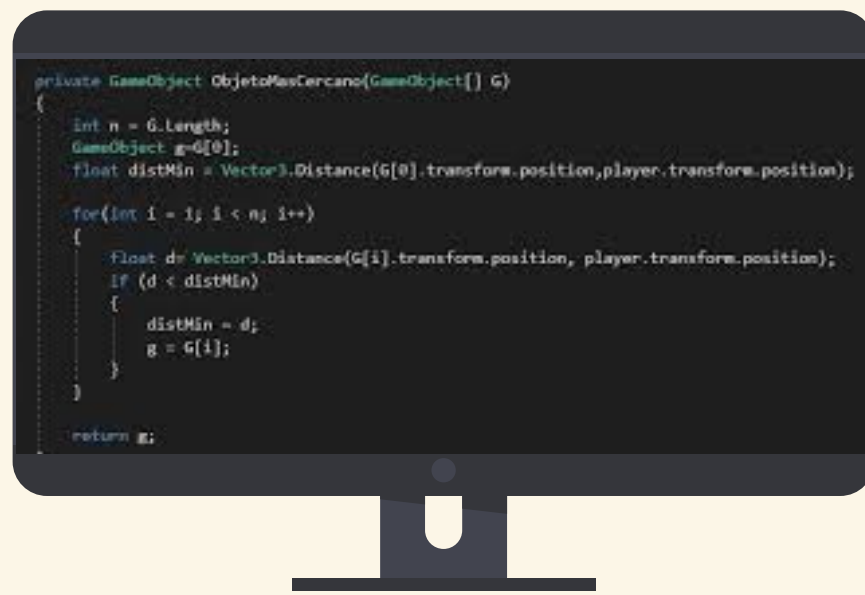
Segmentación manual



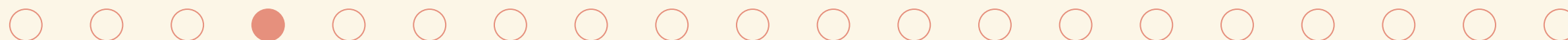
Realizada por radiólogos con experiencia

- ❌ Proceso que consume mucho tiempo
- ❌ Carencia de consistencia y reproducibilidad
- ❌ Sujeto a errores humanos

Algoritmos de aprendizaje profundo



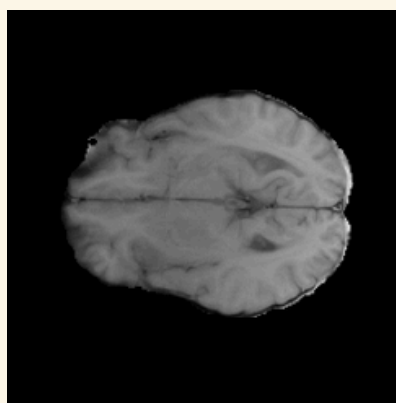
Desafío en diseñar una arquitectura de red neuronal óptima y un programa de entrenamiento adecuado



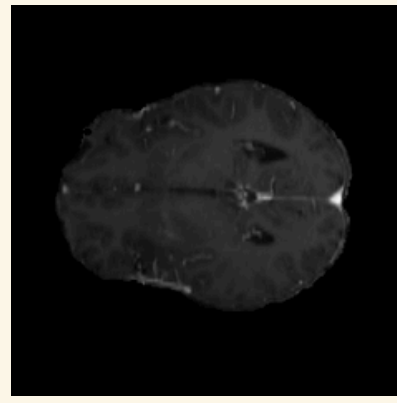
Base de Datos

Se utiliza la base de datos proporcionada por el desafío BraTS21 (Brain Tumor Segmentation Challenge 2021)

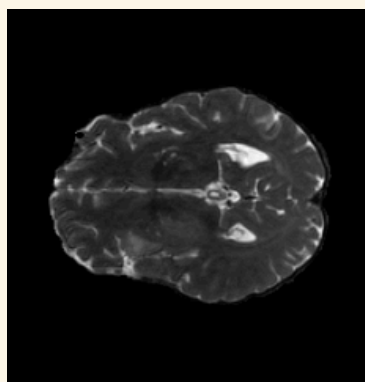
Tipos de imágenes: Imágenes de MRI



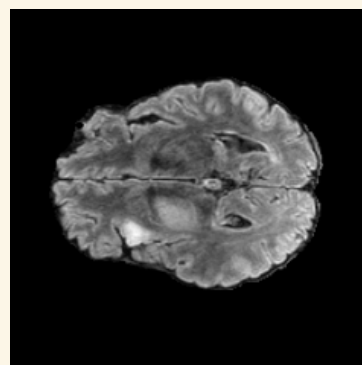
T1: Proporcionan un contraste claro entre la materia gris y blanca



T1Gd: Realzan las áreas con vascularidad de los tumores



T2: Destacan el líquido cerebroespinal y permiten la visualización de edema



FLAIR: Suprime el líquido cerebroespinal para resaltar las lesiones cerebrales

Característica de las imágenes: Almacenadas en *formato NIfTI (.nii)*. Este formato permite manejar volúmenes de datos 3D, preservando la información espacial necesaria para la segmentación precisa.

Características del Conjunto de Datos Resolución:

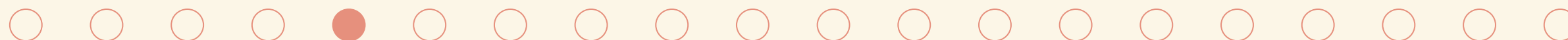
- Preprocesadas con resolución isotrópica de 1mm³, con dimensiones de 240x240x155 voxeles.
- Etiquetas de Segmentación: Las etiquetas incluyen cuatro clases: tumor realzado (ET), tejido edematoso peritumoral (ED), núcleo necrótico del tumor (NCR), y fondo (voxeles que no son parte del tumor).

Anotaciones ✕

Las anotaciones de los tumores han sido realizadas manualmente **por entre uno y cuatro expertos**, lo que asegura una alta calidad en las etiquetas de segmentación.

Conjunto de datos ✕

Entrenamiento: 1,251 casos
Validación: 219 casos
Prueba 570 casos

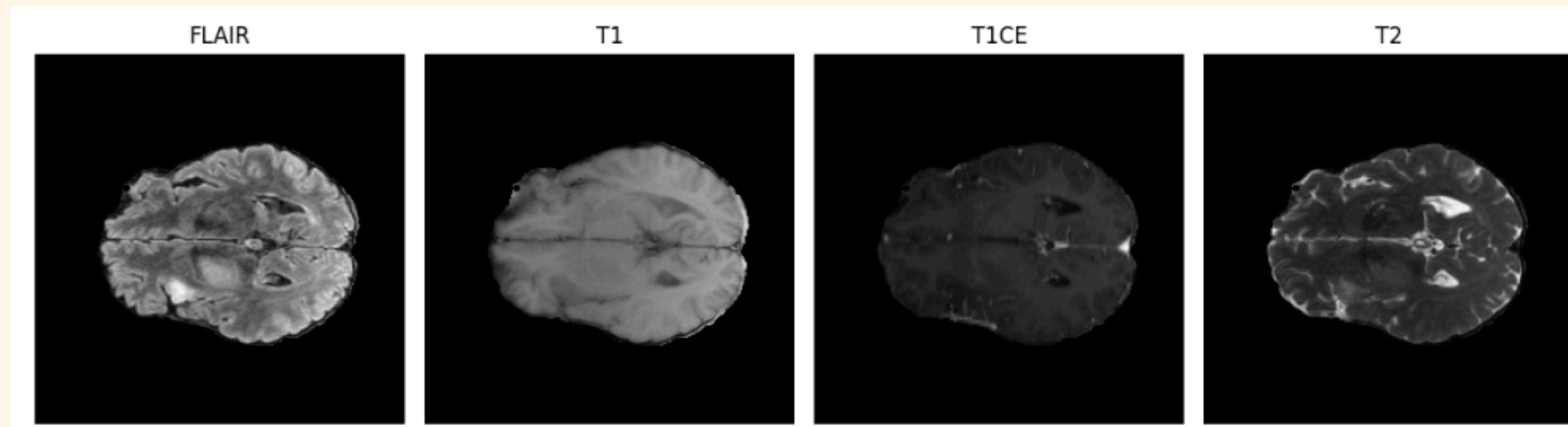


Metodología

Preprocesamiento

Cada ejemplo del conjunto de datos BraTS21 consta de cuatro archivos NIfTI con diferentes modalidades de MRI. Se procedio a apilar las cuatros modalidades de manera que cada ejemplo tenga una forma de (4, 240, 240, 155), de manera que el tensor de entrada tenga el siguiente diseño:

(C, H, W, D) = C-canales, H-altura, W-ancho, D-profundidad



Metodología

Preprocesamiento

- Recorte de Vóxeles de Fondo: Se eliminan los vóxeles de fondo en los bordes del volumen.
- Normalización: Cada canal se normaliza restando la media y dividiendo entre la desviación estándar, sin afectar el fondo.
- Canal Adicional: Se añade un canal de "foreground" con codificación one-hot.
- Almacenamiento: Los volúmenes preprocesados se guardan como arrays de NumPy.

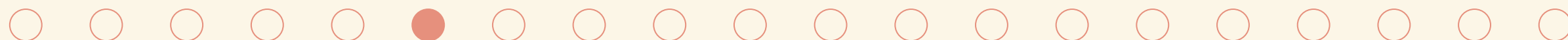


Metodología

Ampliacion de datos

Necesario para evitar el sobreajuste al extender artificialmente un conjunto de datos durante el entrenamiento. Para que el modelo sea más robusto, se utilizaron las siguientes ampliaciones:

- Biased Crop: Recorte de un parche de (5, 128, 128, 128), con un 40% de vóxeles de "foreground".
- Zoom: Se aplica un zoom aleatorio entre 1.0 y 1.4 (15%).
- Flips: Volteo aleatorio a lo largo de los ejes x, y, z (50%).
- Gaussian Noise: Se agrega ruido gaussiano (15%).
- Gaussian Blur: Se aplica desenfoque gaussiano (15%).
- Ajustes de Brillo: Se ajusta el brillo (15%).
- Ajustes de Contraste: Se ajusta el contraste (15%).

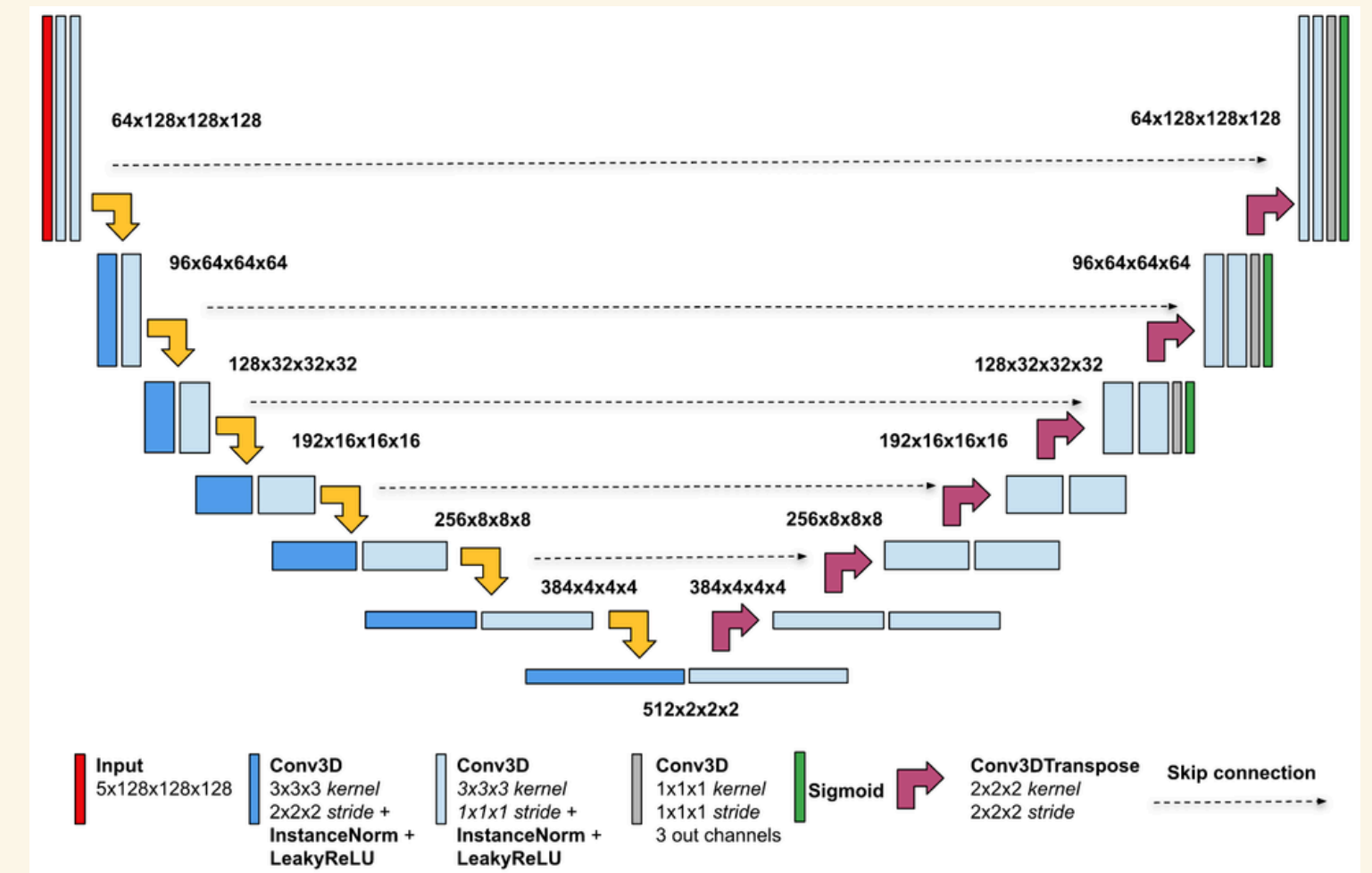


Metodología

Arquitectura de modelo: U-Net

Es una red neuronal diseñada específicamente para tareas de segmentación de imágenes, especialmente en el ámbito médico. Su nombre se debe a su forma distintiva en forma de "U". Se divide en dos partes:

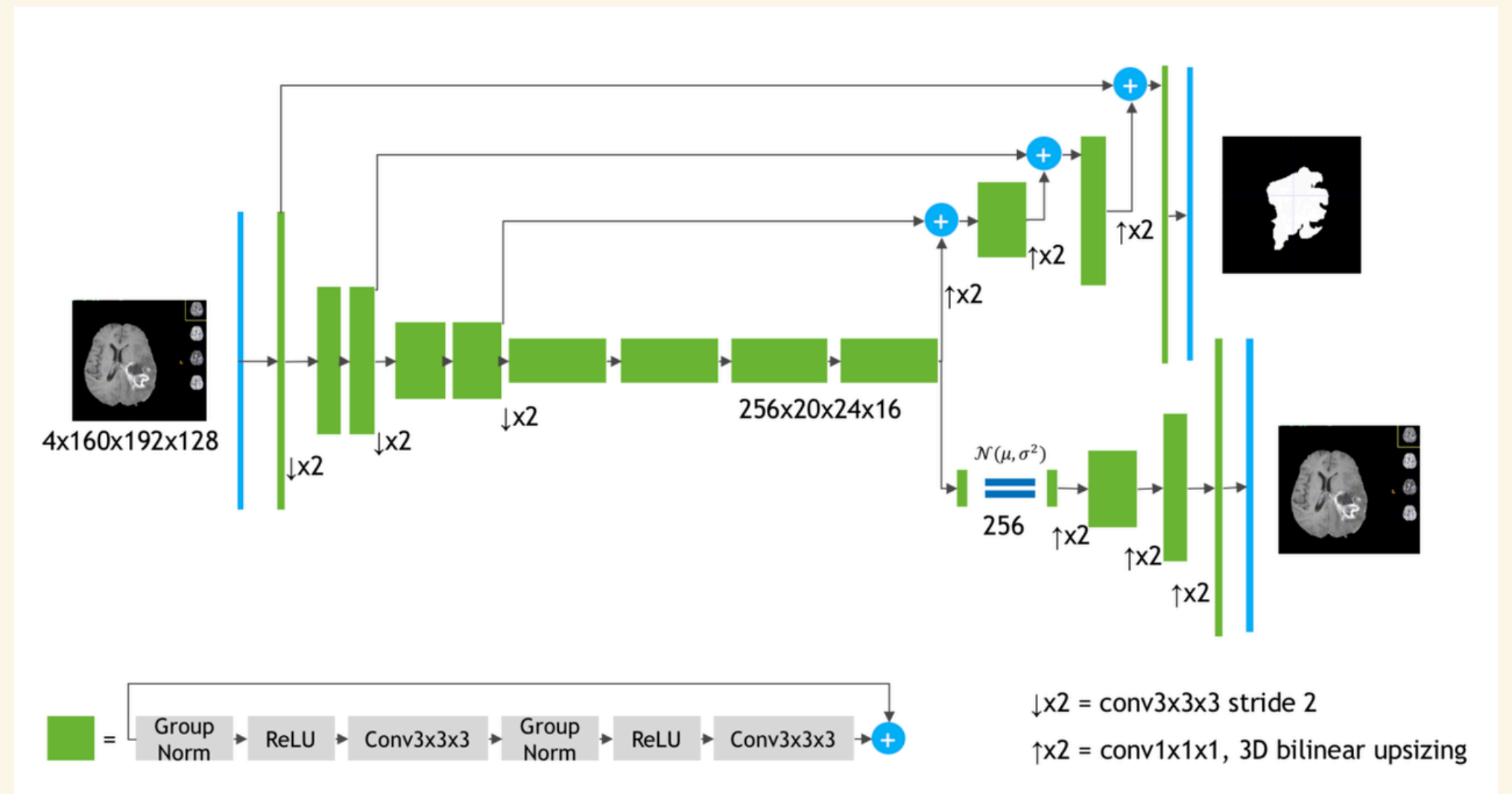
- Codificador: Esta parte de la red se encarga de reducir el tamaño de la imagen original, extrayendo las características más relevantes.
- Decodificador: A partir de la información comprimida del codificador, el decodificador reconstruye una imagen de salida con la misma dimensión que la original, pero ahora cada píxel está clasificado según la categoría a la que pertenece (por ejemplo, tumor, tejido sano, etc.).



Metodología

Arquitectura de modelo: SegResNetVAE

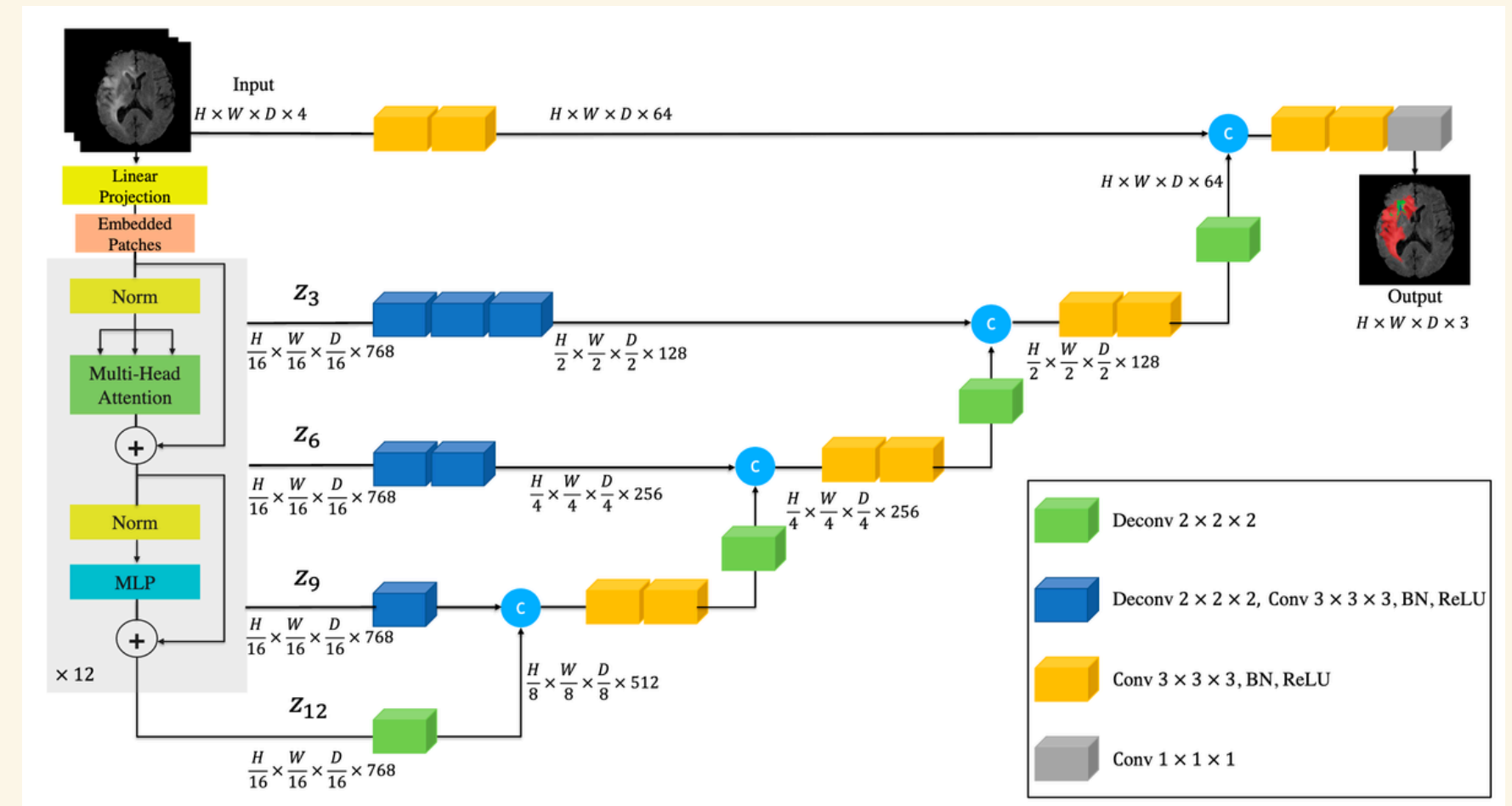
Es una mejora de la arquitectura U-Net, diseñada específicamente para tareas de segmentación de imágenes, especialmente en el ámbito médico. El codificador es similar al U-Net original, pero utiliza bloques residuales para extraer características. El decodificador también es similar al U-Net, pero con una rama adicional de VAE. Esta rama reduce las características del cuello de botella a un espacio de baja dimensión, luego muestrea de una distribución gaussiana y reconstruye la imagen original.



Metodología

Arquitectura de modelo: UNETR

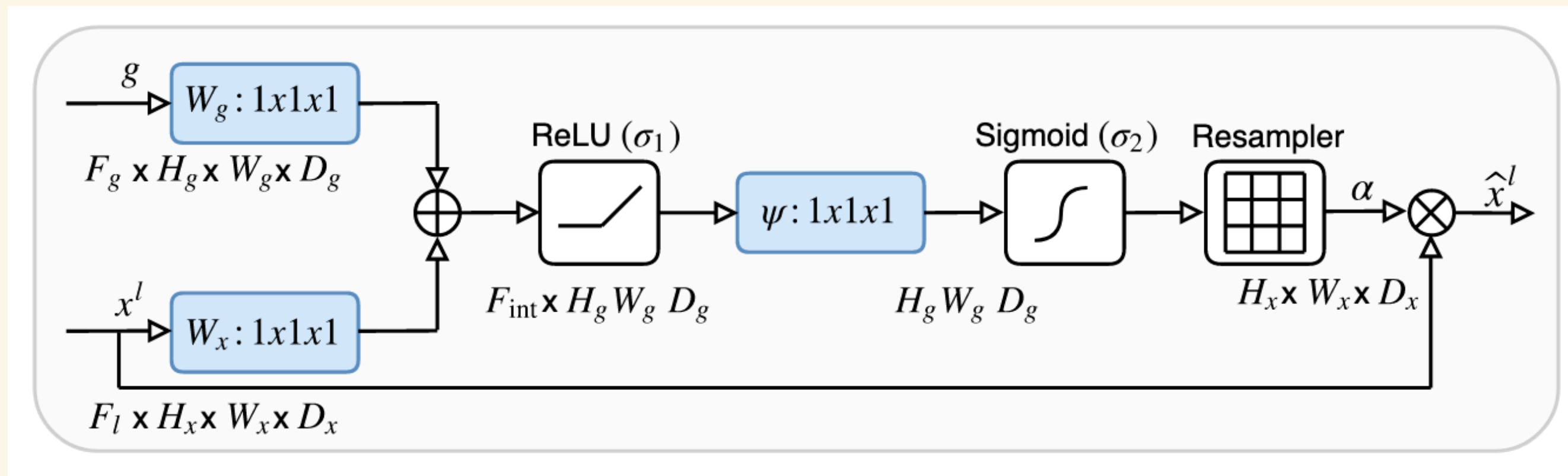
Es una generalización de Vision Transformer (ViT) a las convoluciones 3D: reemplaza las convoluciones 3D en el codificador con autoatención multicabezal. Para convertir un volumen de entrada 3D en una entrada para una autoatención multicabezal, se divide en una secuencia de parches uniformes no superpuestos y se proyecta en un espacio de incrustación utilizando una capa lineal, y se le agrega una incrustación posicional. Luego, dicha entrada se transforma mediante un codificador de autoatención multicabezal.



Metodología

Arquitectura de modelo: Attention U-Net

Extiende la base U-Net añadiendo una compuerta de atención en la parte del decodificador. La compuerta de atención transforma el mapa de características del codificador antes de la concatenación en el bloque del decodificador. Los valores de peso están en el rango (0, 1) y representan el nivel de atención que la red neuronal le presta a un píxel determinado.



Metodología

Inferencia

Durante la inferencia, el volumen de entrada puede tener un tamaño arbitrario, en lugar del tamaño fijo del parche (128, 128, 128) como durante la fase de entrenamiento.

Técnicas de inferencia:

- Sliding window: Para procesar imágenes de cualquier tamaño, se utiliza una ventana deslizante que recorre la imagen, y las predicciones en las regiones superpuestas se promedian.
- Test time augmentations: Se ha creado ocho versiones del volumen de entrada, de modo que cada versión corresponde a uno de los ocho posibles cambios a lo largo de la combinación de los ejes x, y, z.



Metodología

Esta metodología se probó en los conjuntos de validaciones de la validación cruzada de cinco pasos. Se seleccionaron hiperparámetros para obtener la puntuación más alta combinada en todos los pasos. El valor umbral se seleccionó mediante un método de búsqueda en cuadrícula con un paso de 0.05 en el rango (0.3, 0.7). De manera similar, se buscó el número de vóxeles en el rango (0, 100) y se seleccionó maximizando la puntuación en la validación cruzada de cinco pasos.

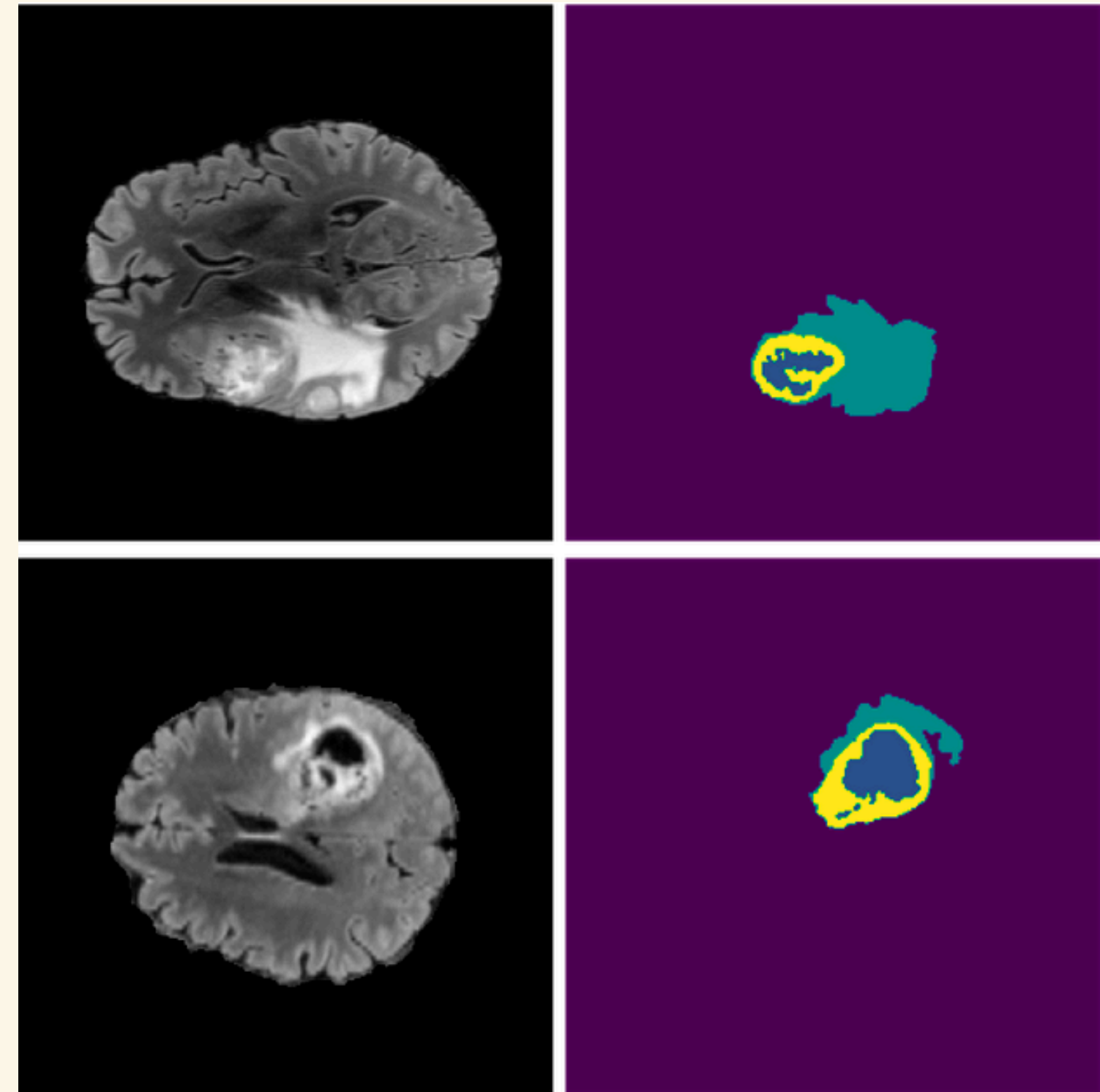


Experimentación

Model	U-Net	UNETR	SegResNetVAE
Fold 0	0.9087	0.9044	0.9086
Fold 1	0.9100	0.8976	0.9090
Fold 2	0.9162	0.9051	0.9140
Fold 3	0.9238	0.9111	0.9219
Fold 4	0.9061	0.8971	0.9053
Mean Dice	0.9130	0.9031	0.9118

Experimentación

Predicciones sobre el conjunto de datos de validación del desafío. En la columna de la izquierda se visualiza la modalidad FLAIR mientras que en la columna de la derecha se visualizan las predicciones del modelo donde el significado de los colores es el siguiente: violeta - fondo, azul - NCR, turquesa - ED, amarillo - ET.



GRACIAS

