

Práctica 3.

Segmentación

Alberto Lardiés, 735976
Devid Dokash, 780131

23 de junio de 2023

Índice

1. Segmentación de imágenes y la red neuronal U-Net	2
2. Resultados	2
2.1. Red neuronal U-net	3
2.2. Thresholding	4
2.3. Otsu	5
2.4. K-means	6

1. Segmentación de imágenes y la red neuronal U-Net

En esta práctica se ha visto la red neuronal convolucional U-net, una red neuronal utilizada para el principal objetivo de la práctica, la segmentación de imágenes.

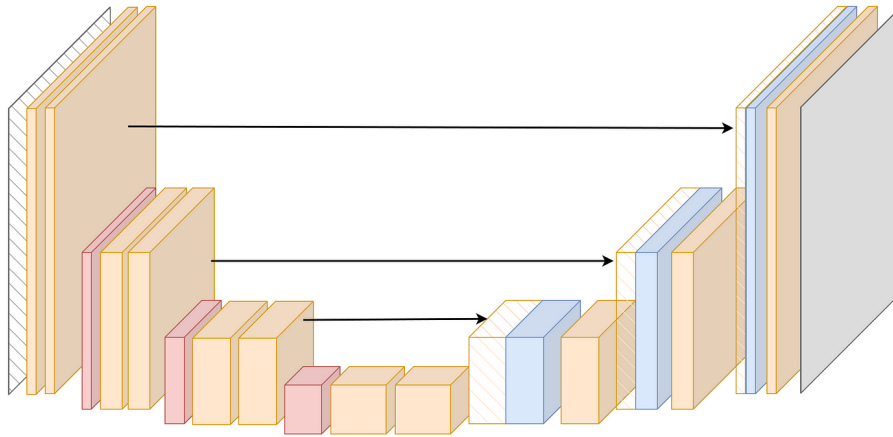


Figura 1: Esquema de la red neuronal U-net.

No se entrará en extenso detalle sobre como se define esta red pero se usará su definición para explicar el siguiente paso en la memoria:

1. **Entrada:** se le pasa como entrada una imagen de un cerebro en escala de grises (luminosidades) como un tensor de sus píxeles.
2. **Capas convolucionales.** Estas capas se encargan de disminuir el tamaño de las imágenes para que sean menos costosas de procesar en pasos posteriores y las representan como estructuras (o mapas) de diferentes características (sea de bordes, texturas o formas) para reconocer posteriormente patrones más complejos y la segmentación en base a estos.
3. **Capas de concatenación.** Estas capas se encargan de unir en cada nivel la capa convolucional de la rama de codificación a la rama de decodificación. Utilizado para poder acceder a resultados y datos obtenidos en diferentes escalas de manera directa en los distintos niveles para las operaciones convenientes (sea recuperar la imagen original o ajustar el modelo para detectar patrones de segmentación más complejos).
4. **Capas convolucionales traspuestas.** Estas capas se encargan de recuperar la información original de la imagen y devolverá a su estado original, recuperando en cada nivel de convolución un nivel de detalle mayor para poder detectar los patrones más complejos y ajustar mejor el modelo.
5. **Capa de convolución de salida.** La capa final de la red, con el modelo ya ajustado, identificará la probabilidad de que un píxel sea de cada clase, seleccionando la clase con mayor probabilidad y finalmente, consiguiendo la segmentación de las imágenes.

2. Resultados

En este apartado se profundiza en los resultados de la arquitectura U-net, evaluados mediante la métrica Dice Similarity Coefficient y comparados con un algoritmo tradicional, el thresholding. A continuación las tando de muestras utilizadas para la evaluación:

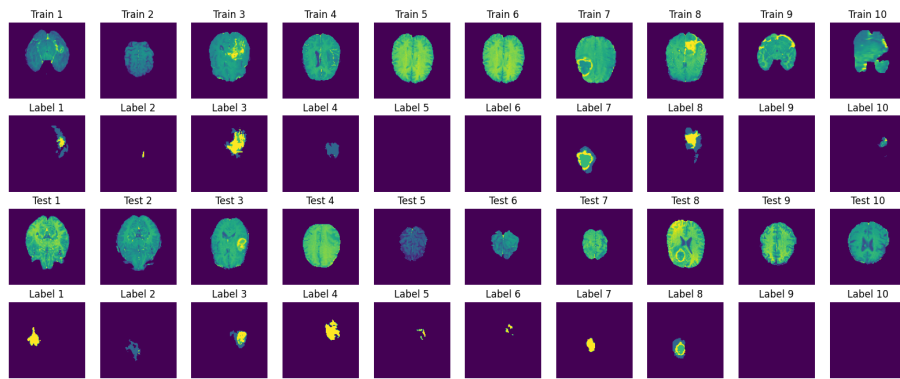


Figura 2: Esquema de la red neuronal U-net.

2.1. Red neuronal U-net

Se puede observar como la red neuronal obtiene resultados bastante buenos, aunque en algunas de las muestras pueda fallar más o menos y la segmentación difiera ligeramente de las segmentaciones originales.

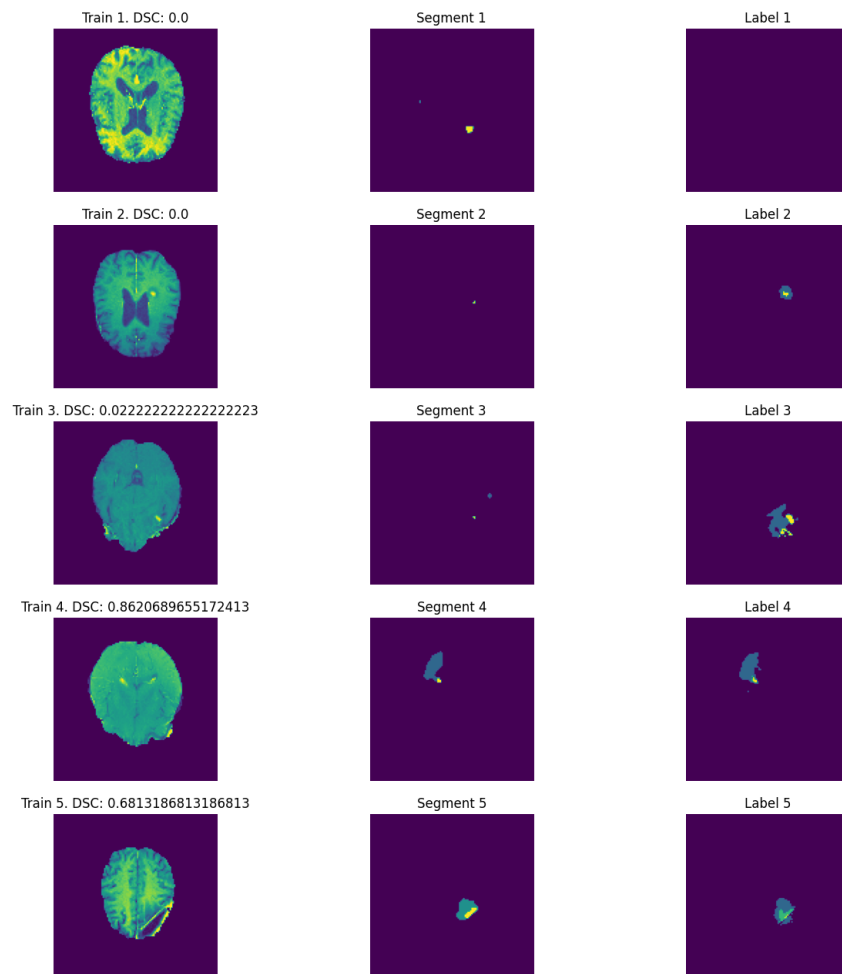


Figura 3: Esquema de la red neuronal U-net.

2.2. Thresholding

Se ha seleccionado como algoritmo de comparación el thresholding por su sencillez. La principal diferencia del thresholding es la manera que determina la segmentación de las imágenes.

Mientras la red neuronal trata las imágenes y detecta patrones para ajustar al modelo, el thresholding ignora totalmente las características de las imágenes y de sus atributos y depende totalmente de la intensidad de las muestras.

En otras palabras, mientras que la red neuronal se basa en las características de sus muestras para encontrar patrones y relaciones que identifiquen las diferentes segmentaciones de la imagen, un algoritmo tradicional como el thresholding se basa en la intensidad de los píxeles, ignorando totalmente la relación entre los mismos o los patrones que pudiese haber.

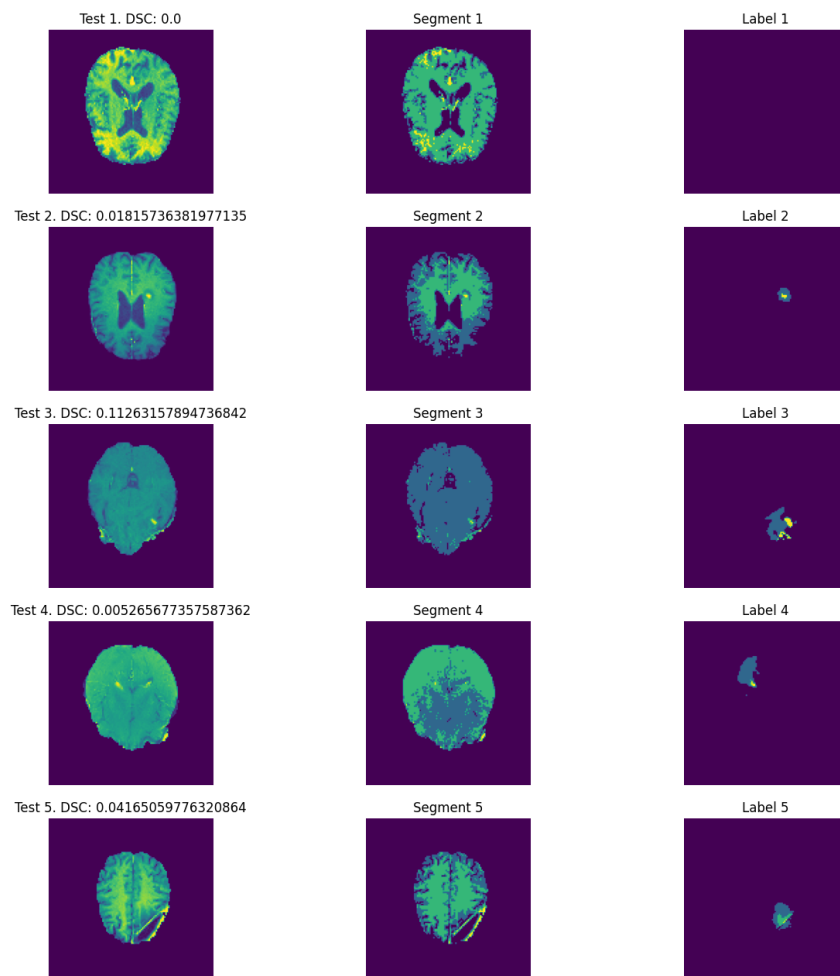


Figura 4: Resultados de la segmentación con Thresholding.

2.3. Otsu

Segundo algoritmo tradicional para la segmentación de las imágenes. El algoritmo de Otsu (en este caso, multi-otsu), se basa en encontrar el mejor umbral para aplicar el thresholding de manera que maximice la varianza inter-clase (en este caso cuatro clases):

1. Se aplica un algoritmo recursivo. Se haya el mejor umbral para separar de manera binaria el fondo y los objetos de la imagen.
2. Una vez calculada, se vuelve a llamar al algoritmo, y el rango de intensidades esta vez se realiza desde el mínimo hasta el mejor umbral por una parte y desde el mejor umbral hasta la máxima intensidad por otra.
3. Se repite el proceso hasta que el histograma de la imagen queda dividida en tantas clases, en este caso, cuatro (tres umbrales).
4. Los resultados no son bastante buenos, y no se sabe si es por que si el algoritmo en si no funciona o las imágenes no es posible segmentarlas bien con diferentes umbrales.

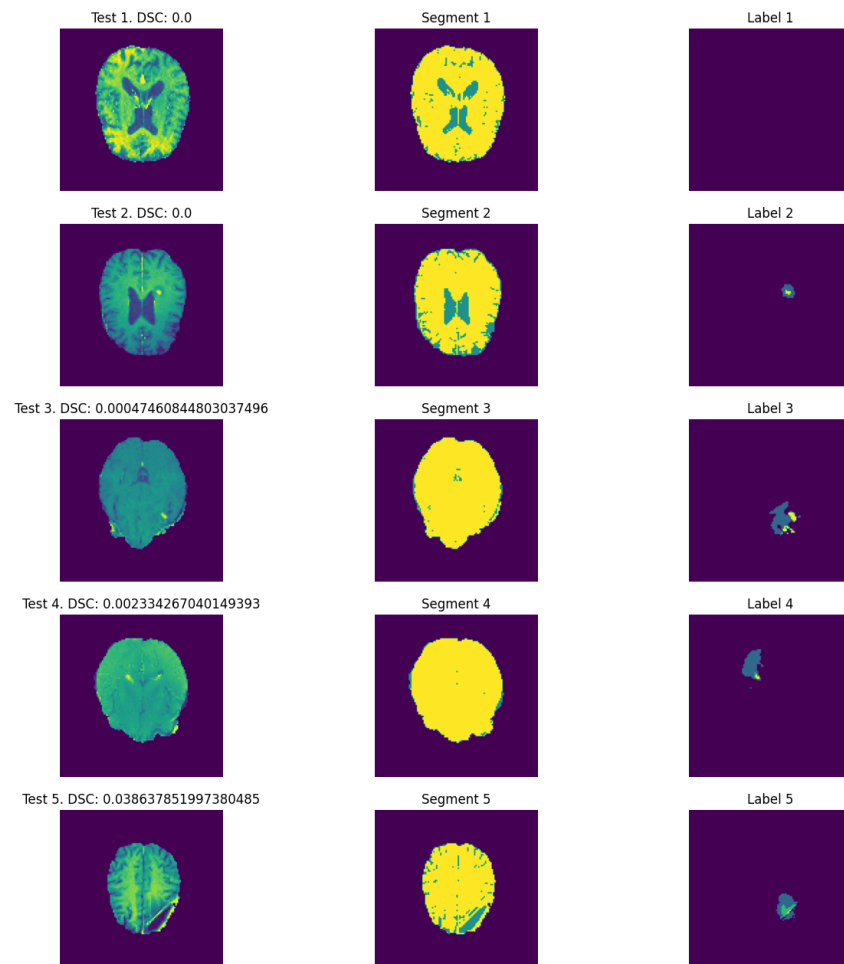


Figura 5: Resultados de la segmentación con Otsu.

2.4. K-means

El tercer y último algoritmo tradicional para la segmentación de las imágenes. El algoritmo de clustering de K-means se basa en agrupar los diferentes valores, las intensidades en este caso, en clusters bajo una cierta heurística (la distancia euclídea en este caso):

1. Se proponen cuatro clusters, uno por cada clase.
2. Los resultados del algoritmo de K-means generalmente saldrán mal, esto es debido a la selección de los centroides iniciales y el como se van conformando los clusters.
3. El problema de este algoritmo es que dependiendo de la posición de los centroides, es posible que haya clusters que predominen frente a otros y no necesariamente sean correctos (debido a la posición de los mismos frente a los distintos valores, tengan más predominancia frente a otros clusters y puede que realmente no pertenezcan a este cluster).

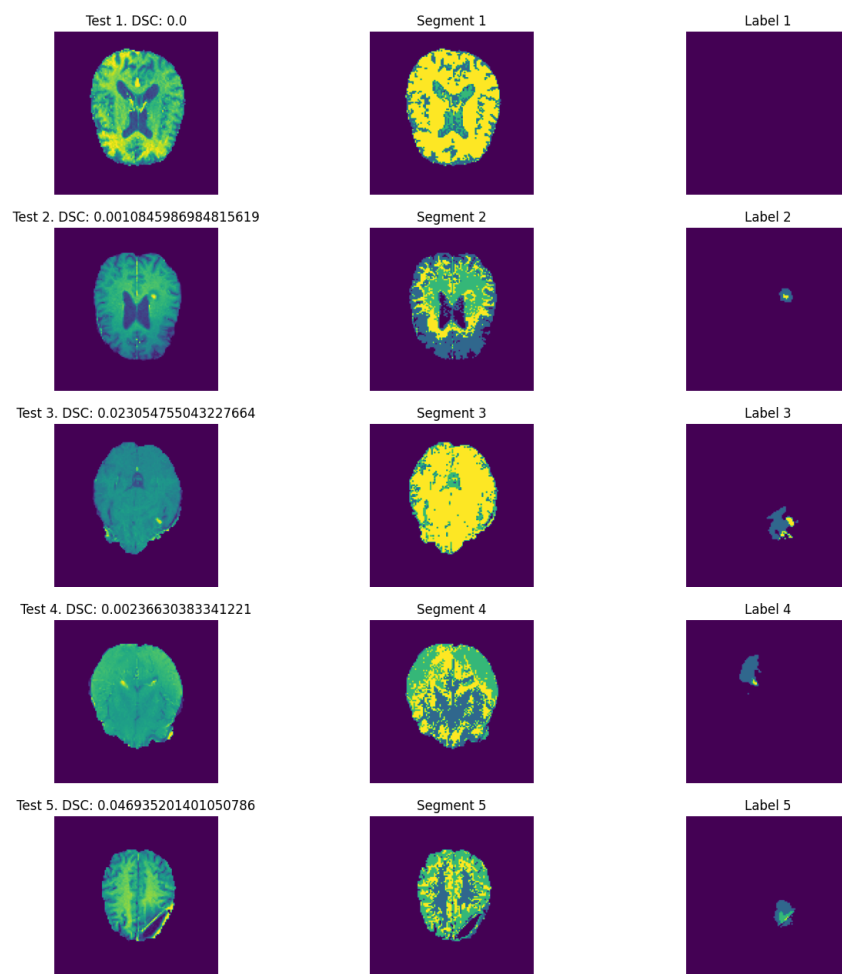


Figura 6: Resultados de la segmentación con K-means.

4. Para sesgar un poco este problema, se ha forzado a que uno de los centroides tenga como valor inicial el 0, forzando a que el cluster conforme el fondo. Para los demás centroides, se ha forzado que las intensidades seleccionadas de manera aleatoria estén en la parte central de la imagen y sean distintas de 0.
5. Se podría optar como solución forzar diferentes niveles de intensidad a cada centroide, es decir, que para cada centroide, aplicar un rango específico de intensidades (centroide 1 = 0,

centroide 2 = 120, centroide 3 = 180, centroide 4 = 220...), pero aunque pueda dar mejores resultados que todos los demás métodos mencionados, se acentuaría el problema de depender de las intensidades. La decisión dependerá radicaré en las necesidades del problema.

6. K-means sufre del mismo problema que el thresholding y Otsu, y es que dependen únicamente de las intensidades y no tienen en cuenta otras características de la imagen que pudiesen identificar mejor la segmentación (no porque tenga una alta intensidad ciertas partes del cerebro significa que este tenga un problema).