

# Segmentación Semántica de lesiones de Ictus Isquémico en imágenes de Resonancia Magnética

T. Vargas Rodríguez, S. Gutiérrez Díaz, D. Mancheño Castaño. J.L. Pareja Mora

<sup>1</sup> Departamento de Matemática Aplicada, Ciencia e Ingeniería de los Materiales y Tecnología Electrónica, Universidad Rey Juan Carlos, Madrid, España, {t.vargas.2020, s.gutierrezd.2020,d.mancheno.2021, jl.pareja.2020}@alumnos.urjc.es

## Resumen

*Este estudio se centra en la aplicación de la red neuronal U-Net para la segmentación semántica de lesiones de ictus isquémico en imágenes de resonancia magnética ponderada en difusión (DWI, Difusión Weighted Image y su mapa de atenuación, ADC). Para ello, hemos empleado una base de datos balanceada con 30 imágenes de resonancia magnética DWI en formato NIFTI, obtenida de dos centros diferentes y con máquinas distintas. El preprocesamiento inicial involucró la normalización del tamaño y la intensidad de los píxeles y la eliminación de cortes con más de un 98% de píxeles negros, pues no nos servirían para entrenar nuestra red. La red U-Net ha sido seleccionada como el modelo para realizar esta segmentación. Una vez entrenada y evaluada la red, la accuracy obtenida fue baja, pero el modelo demostró una gran precisión, un alto valor de ROC-AUC y de coeficiente IOU (Intersection Over Union) y una baja tasa de pérdida, demostrando una capacidad efectiva para distinguir entre tejidos afectados y no afectados por el ictus.*

## 1. Introducción

El ictus [1], o accidente cerebrovascular, ocurre cuando el flujo de sangre al cerebro es reducido o interrumpido, privando al tejido cerebral de oxígeno y nutrientes esenciales, y provocando su muerte si no se trata a tiempo. La detección en las fases agudas y subagudas de estas lesiones es crucial para su seguimiento clínico y para definir estrategias terapéuticas y de rehabilitación óptimas, maximizando la recuperación y minimizando las secuelas en el paciente. Para la detección del infarto cerebral, hemos analizado en profundidad 1500 imágenes diferentes de resonancia magnética ponderada en difusión (DWI) de 30 pacientes. Sin embargo, la interpretación de estas imágenes puede ser compleja, tediosa y longeva, además de estar sujeta a variabilidad entre operadores. Consecuentemente, hemos decidido implementar técnicas de segmentación semántica automatizada basadas en la red neuronal U-Net. Esta es una arquitectura diseñada específicamente para la segmentación semántica de imágenes biomédicas de manera que su uso nos proporciona una mayor precisión y eficiencia en la detección de lesiones de ictus.

## 2. Materiales y Métodos

### 2.1. Base de Datos

La base de datos utilizada ha sido obtenida del sitio web “Grand Challenge” [2], concretamente del challenge “Ischemic Stroke Lesion Segmentation Challenge - ISLES'22” [3]. Este challenge propone un desafío de segmentación semántica de lesiones de ictus. La base de datos que proporciona para el entrenamiento del modelo contiene imágenes de resonancia magnética de 250 pacientes diferentes -personas mayores de edad que habían recibido

imágenes por resonancia magnética del cerebro por un accidente cerebrovascular previamente diagnosticado o sospechado- con alta variabilidad en tamaño, cantidad y ubicación de las lesiones. Presenta tres modalidades de imagen distintas: *Difusión Weighted Image* o DWI; su mapa de atenuación correspondiente, o ADC; y la modalidad *Fluid Attenuated Inversion Recovery* o FLAIR. Además, esta base de datos incluye la máscara de segmentación o *Ground Truth*, las cuales se corresponden con segmentaciones realizadas a mano por diferentes expertos en la detección de estas lesiones. También incluye a 5 pacientes que a los que no se les detectó ninguna lesión.

La naturaleza multicéntrica de la base de datos aporta una significativa diversidad, además de incorporar imágenes adquiridas con diferentes máquinas (3T Philips, 3T Siemens y 1.5T Siemens) y protocolos de dos centros distintos. Esta variedad multicéntrica es muy importante para poder crear un algoritmo robusto y generalizable frente a las distintas calibraciones de las máquinas. [4-5]

A pesar de disponer de 250 pacientes, debido a limitaciones de la memoria RAM, no hemos podido entrenar nuestra red incluyendo a todos los pacientes. Consecuentemente, hemos seleccionado cuidadosamente una muestra de 30 pacientes, 15 de cada centro, para asegurar una representación equitativa de las adquisiciones y una base de datos final balanceada. Además, el modelo de segmentación final se ha basado únicamente en las modalidades DWI y ADC por la misma razón. Todo esto podría resultar en una limitación de la capacidad del modelo para aprender de forma más generalizada y de aprender a segmentar otras modalidades de resonancia magnética.

La elección DWI como modalidad a segmentar se debe a su relevancia clínica en la detección temprana de lesiones de ictus [6]. Esta modalidad se basa en la detección en vivo del movimiento de moléculas de agua, la cual puede estar restringida en algunas áreas de tejido por la falta de flujo sanguíneo. Estas imágenes son de fácil y sencilla interpretación, y son capaces de mostrar alteraciones isquémicas en el cerebro antes de que aparezcan otros cambios evidentes en tomografía computarizada (TC) o secuencias convencionales de resonancia magnética (RM). La información obtenida de esta modalidad se ve complementada con la introducción de los mapas de atenuación (ADC). Su inclusión en el estudio es fundamental para mejorar la detección y caracterización de lesiones de ictus isquémico. Los mapas ADC, al proporcionar una cuantificación precisa del movimiento de las moléculas de agua, permiten diferenciar de manera más efectiva entre el tejido isquémico real y los artefactos presentes en las

imágenes DWI. Esta capacidad de diferenciación aumenta la precisión en la identificación de las alteraciones isquémicas, lo que resulta en un diagnóstico más fiable y una comprensión más completa del daño cerebral.

Por último, es importante señalar que las imágenes proporcionadas proceden de pacientes completamente anónimos de forma irreversible, y están en formato NIFTI. No han sido registradas entre sí, de manera que no se ha realizado un alineamiento inicial entre imágenes de diferentes sujetos. Además, a las imágenes se le han eliminado el cráneo, mostrando únicamente el tejido cerebral, reduciendo así la complejidad de la segmentación.

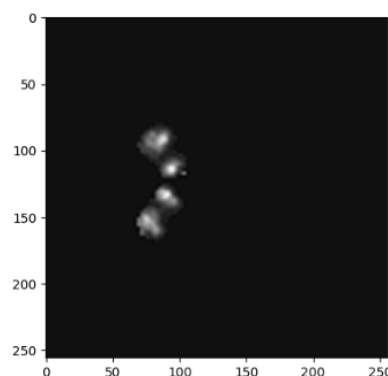
## 2.2. Preprocesado

Se ha llevado a cabo un meticuloso proceso de normalización y filtrado de imágenes de la base de datos, con el objetivo de optimizar la homogeneidad y calidad del conjunto de datos. Este enfoque se ha estructurado en dos fases fundamentales: la normalización por tamaño e intensidad, y el descarte selectivo de imágenes basado en el percentil 98.

Para la normalización por tamaño hemos redimensionado las imágenes al mismo tamaño (256x256) mediante una interpolación cúbica. Este proceso es esencial para mitigar posibles distorsiones que puedan surgir debido a diferencias en las resoluciones originales de las imágenes. La consistencia en las dimensiones facilita el análisis comparativo y contribuye a la robustez de los resultados obtenidos. En segundo lugar, se ha llevado a cabo una normalización por intensidad, garantizando la homogeneidad en los niveles de grises y contraste en todas las imágenes. La normalización por intensidad resulta esencial para evitar sesgos en los análisis posteriores y asegurar la precisión de los resultados.

Por otro lado, hemos realizado la binarización de las máscaras asociadas a las imágenes. Este procedimiento ha sido implementado con el objetivo de simplificar la representación de la información contenida en las máscaras, transformando las imágenes a un formato binario donde los píxeles adquieren valores de blanco o negro (lesión o no). Este proceso es esencial ya que tenemos que distinguir claramente las regiones de interés y el fondo.

Adicionalmente, hemos implementado un criterio de filtrado basado en el percentil 98 mediante el cual desechamos las imágenes que contienen un 98% de píxeles sin información (imágenes con una alta cantidad de píxeles negros). La exclusión de imágenes con información limitada o poco representativa fortalece la integridad de los análisis y asegura la fiabilidad de los resultados obtenidos durante la investigación. La Figura 1 muestra la cantidad mínima de información que debe tener un corte para considerarlo en entrenamiento.



**Figura 1.** Cantidad mínima visual de píxeles diferentes de negro para considerar la imagen para el entrenamiento

## 2.3. U-Net [7-8]

Para lograr una segmentación semántica precisa de lesiones de ictus, hemos decidido usar la red convolucional U-Net. Esta red tiene gran capacidad para aprender eficazmente de un número limitado de imágenes de entrada, aumentando artificialmente la variabilidad y el volumen del conjunto de entrenamiento. Además, es una red relativamente rápida en comparación con otras redes profundas, y puede generar segmentaciones en tiempo casi real, siendo esto muy importante para el diagnóstico precoz de ictus.

La arquitectura de la red U-Net consta de dos caminos: un camino contractivo, el cual permite capturar información contextual y de características, y un camino expansivo, que permite la localización precisa en la segmentación.

Este modelo se caracteriza por su simetría, conformando una estructura en forma de “U”. La principal ventaja de la U-Net proviene de la presencia de conexiones de salto, que permiten el intercambio de información mediante la concatenación de la salida de cada capa del camino contractivo con su capa simétrica en el camino expansivo. Estas conexiones son esenciales para ayudar a recuperar la información de localización espacial que se pierde en el camino contractivo.

Nuestro modelo está compuesto por unos bloques llamados módulos de convolución. Estos se basan en la concatenación de dos convoluciones con filtros de tamaño 3x3, configuradas con *padding* constante y función de activación *ReLU*. El relleno constante se emplea para mantener las dimensiones de las imágenes y asegurar la precisión en la localización espacial.

En total, nuestro modelo cuenta con 23 convoluciones con filtros de tamaño 3x3.

El camino contractivo incluye cuatro módulos de convolución, cada uno seguido de una capa de *max pooling* con tamaño 2x2 y con *stride* = 2 para reducir a la mitad la dimensión de la imagen (*downsampling*). A medida que se avanza en la red, la representación se profundiza con más filtros, permitiendo a la red capturar contextos más complejos.

De igual manera, el camino expansivo está formado por cuatro módulos de convolución, cada uno precedido de *max*

*unpooling* mediante convolución transpuesta, para aumentar la dimensión de las imágenes (*upsampling*). A continuación, se realiza la concatenación del mapa de características correspondiente. El objetivo de esta capa es aumentar de forma progresiva la resolución de las imágenes.

En cada etapa de *max pooling* y *up pooling* se ha integrado un *dropout* con valor 0.3. Esto hace que en cada capa se “desactiven” el 30% de las neuronas, evitando así que el modelo aprenda de ellas y se sobreajuste. Además, se ha empleado la técnica de *Early Stopping* en el entrenamiento para evitar el sobreajuste de la red. Este método se basa en la monitorización de las figuras de mérito del modelo y detiene el entrenamiento cuando no se observan mejoras en el valor de las métricas evaluadas para el conjunto de validación.

El modelo construido permite una segmentación detallada y precisa, crucial para la identificación y diagnóstico de lesiones de ictus. La estructura de la U-Net se presenta en la Figura 2.

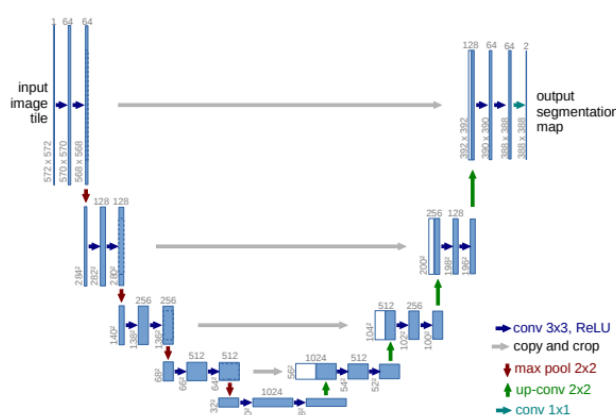


Figura 2. U-Net

### 3. Resultados

En la evaluación del modelo de segmentación U-Net hemos obtenido resultados significativos:

El modelo alcanzó una **tasa de pérdida** (loss) de solo 0.556. Esto nos indica que la red fue capaz de realizar predicciones cercanas a las segmentaciones de referencia (*Ground Truth*) que nos proporciona la base de datos. Un ejemplo de esto se puede ver en la Figura 3.

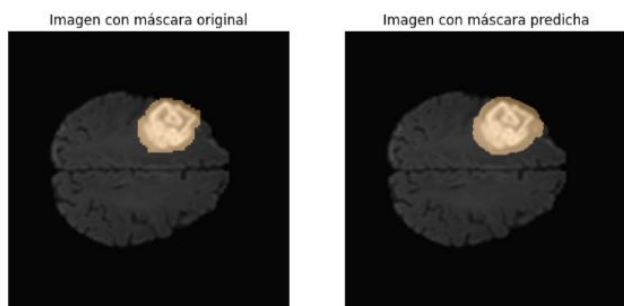


Figura 3. A la izquierda, máscara de segmentación *Ground Truth*. A la derecha, segmentación predicha por la U-Net

Para realizar una comparativa entre el volumen real y el volumen predicho de lesión se ha analizado el coeficiente de Intersección sobre Unión (**IOU**). Este coeficiente nos indica que la superposición de las áreas segmentadas por la U-Net y la máscara *Ground Truth*. Así, se alcanzó un valor del 80.91%, lo cual muestra un alto grado de precisión en la segmentación realizada por el modelo. Con esta métrica, se concluye que la red no solo está identificando correctamente las áreas relevantes, sino que también evita segmentar erróneamente áreas que no son lesiones.

La **tasa de acierto** (accuracy) obtenida fue del 25.126%. Este valor es bastante bajo, pero es importante considerar la complejidad de la tarea. La tasa de acierto no solo se ve afectada por la clasificación correcta de los píxeles, sino también de la precisión en la delimitación del área afectada. En la figura 4 se puede observar que la delimitación de la segmentación realizada por la red no coincide al completo con el *Ground Truth*, lo cual influye en el valor de la tasa de acierto.

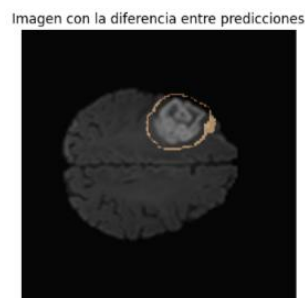


Figura 4. Diferencia entre las máscaras *Ground Truth* y la máscara predicha

Uno de los aspectos más destacados es el alto valor del área bajo la curva ROC (**ROC-AUC**). Este valor es del 99.096%, y nos indica que el modelo tiene una excelente capacidad para discriminar las clases objetivo, en este caso, áreas afectadas y no afectadas por ictus.

Por último, la **precisión** (precision) alcanzó un valor del 91.511%, lo que refleja una gran capacidad para identificar de manera correcta las lesiones de ictus en las imágenes. Esto tiene gran relevancia clínica, pues la identificación precisa de las lesiones es fundamental para el diagnóstico correcto y la planificación del tratamiento.

En conjunto, los resultados demuestran que hemos creado un modelo bastante efectivo en la segmentación de lesiones de ictus en imágenes DWI y ADC. La combinación de una baja tasa de pérdida, además de un alto valor de ROC-AUC y precisión indican que el modelo es bastante robusto. No obstante, una tasa de acierto tan baja señala que todavía hay áreas en las que nuestro modelo puede mejorar.

### 4. Discusión

A pesar de que el modelo creado con la red U-Net ha demostrado una gran capacidad de segmentación semántica de múltiples lesiones de ictus en imágenes de resonancia

magnética, hemos encontrado varias limitaciones y áreas de mejora que deben ser consideradas para trabajos futuros.

Una de las principales limitaciones del estudio fue el uso de un conjunto de datos relativamente pequeño, limitado a imágenes de resonancia magnética de únicamente 30 pacientes (de los 250 que proporcionaba inicialmente la base de datos). Esto se debe a la falta de capacidad computacional para procesar tantas imágenes, lo cual puede haber afectado a la capacidad del modelo para aprender más características y adaptarse a la variabilidad de las imágenes.

De manera similar, tampoco se contaba con suficiente capacidad computacional para incluir la modalidad FLAIR en el entrenamiento de la red neuronal. Esta modalidad puede proporcionar información valiosa sobre el tejido cerebral y mejorar la identificación y delimitación de lesiones del modelo.

En las fases iniciales del proyecto exploramos la posibilidad de utilizar la red 3D-Unet [9-10] para mejorar la segmentación de volúmenes. No obstante, sólo pudimos probarla para imágenes de resonancia magnética en la modalidad DWI, obteniendo unas segmentaciones muy deficientes. Dejamos para investigaciones futuras el uso de esta red incorporando las tres modalidades existentes (DWI, ADC y FLAIR), pues creemos que la combinación de estas modalidades podría mejorar significativamente la precisión, aprovechando la riqueza de la información tridimensional.

Para mejorar las métricas obtenidas, sería conveniente usar un hardware más potente para poder manejar un mayor volumen de datos, permitiendo así la inclusión de un mayor número de imágenes con información más diversa y la inclusión de la modalidad FLAIR. Esto ayudaría a superar las limitaciones actuales y potenciaría la capacidad del modelo para generalizar mejor en un conjunto de datos más amplio y variado.

## 5. Conclusión

Este estudio exploró la aplicación de la red neuronal U-Net para la segmentación semántica de lesiones de ictus en imágenes de resonancia magnética ponderadas en difusión (DWI) y sus mapas de atenuación (ADC). A pesar de las limitaciones en cuanto al tamaño del conjunto de datos usado y a la exclusión de la modalidad FLAIR en la etapa de entrenamiento, hemos conseguido unos resultados significativos y reveladores.

La red U-Net ha demostrado una gran capacidad para segmentar lesiones. Esto se ve reflejado en el alto valor de ROC-AUC y de precisión, y en el bajo porcentaje de tasa de pérdida obtenido. Por contraposición, la baja tasa de acierto destaca la complejidad de la tarea y señala la necesidad de mejoras en el modelo y entrenamiento. De esta manera, un conjunto de imágenes más extenso mejoraría la generalización y eficacia del modelo.

En conclusión, los resultados obtenidos son prometedores y abre caminos para avances futuros en la aplicación de técnicas de aprendizaje profundo en el análisis de imágenes médicas.

## Referencias

- [1] Ictus-ACV síntomas y tratamiento | Clínica Universidad Navarra. (2023, 23 septiembre). <https://www.cun.es/enfermedades-tratamientos/enfermedades/ictus#:~:text=El%20accidente%20cerebrovascular%20o%20ictus,de%20ox%C3%ADgeno%20y%20nutrientes%20esenciales> (consultada: Diciembre 2023)
- [2] Grand Challenge. (s. f.). [grand-challenge.org](https://grand-challenge.org/). <https://grand-challenge.org/> (consultada: Diciembre 2023)
- [3] Ischemic Stroke Lesion Segmentation Challenge - Grand Challenge. (s. f.). [grand-challenge.org](https://isles22.grand-challenge.org/home/). <https://isles22.grand-challenge.org/home/> (consultada: Diciembre 2023)
- [4] Ischemic Stroke Lesion Segmentation Challenge 2022: Acute, Sub-acute and Chronic Stroke Infarct Segmentation. (2022). Zenodo. <https://doi.org/10.5281/zenodo.6362388> (consultada: Diciembre 2023)
- [5] Petzsche, M. R. H., De La Rosa, E., Hanning, U., Wiest, R., Valenzuela, W., Reyes, M., Meyer, M. I., Liew, S., Kofler, F., Ezhov, I., Robben, D., Hutton, A., Friedrich, T., Zarth, T., Bürkle, J., Baran, A., Menze, B., Broocks, G., Meyer, L., . . . Kirschke, J. S. (2022). ISLES 2022: A multi-center Magnetic Resonance Imaging stroke lesion Segmentation dataset. Scientific Data, 9(1). <https://doi.org/10.1038/s41597-022-01875-5> (consultada: Diciembre 2023)
- [6] S, P. S. (2009b). DIFUSIÓN POR RESONANCIA MAGNÉTICA: BASES y APLICACIONES ONCOLÓGICAS EN ÓRGANOS EXTRACRANEANOS. Revista Chilena de Radiología, 15. <https://doi.org/10.4067/s0717-93082009000400004> (consultada: Diciembre 2023)
- [7] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-NET: Convolutional Networks for Biomedical Image Segmentation. arXiv (Cornell University). <https://doi.org/10.48550/arxiv.1505.04597> (consultada: Diciembre 2023)
- [8] Maynard-Reid, M. (2023, 23 marzo). U-Net image segmentation in Keras - PyImageSearch. PyImageSearch. <https://pyimagesearch.com/2022/02/21/u-net-image-segmentation-in-keras/> (consultada: Diciembre 2023)
- [9] Segmentation Models Python API — Segmentation Models 0.1.2 documentation. (s. f.). <https://segmentation-models.readthedocs.io/en/latest/api.html#unet> (consultada: Diciembre 2023)
- [10] Bnsreenu. (s. f.). Python\_for\_microscopists/215\_3D\_Unet.ipynb at master · bnsreenu/python\_for\_microscopists. GitHub. [https://github.com/bnsreenu/python\\_for\\_microscopists/blob/master/215\\_3D\\_Unet.ipynb](https://github.com/bnsreenu/python_for_microscopists/blob/master/215_3D_Unet.ipynb) (consultada: Enero 2024)