



Segmentación de la Materia Blanca Cerebral por Medio de Aprendizaje Automático

Posadas Gámez Eduardo Daniel¹, Tovar Arriaga Saúl¹ ✉, Ortiz Feregrino Rafael¹,
Pasaye Alcaraz Erick Humberto²

¹Universidad Autónoma de Querétaro, Facultad de Ingeniería, Querétaro, México,

²Universidad Nacional Autónoma de México, Instituto de Neurobiología, Querétaro, México.
✉ saul.tovar@uaq.mx

Resumen

La materia blanca tiene un papel importante en la estructura cerebral, y una lesión o degeneración en ésta área, puede causar daños severos a todo el sistema cerebral, por lo tanto, la segmentación automática de imágenes médicas cerebrales es un proceso importante al utilizar software de apoyo para la detección de lesiones, neurodegradación o en la evaluación de neurocirugías. En esta investigación, se realizará un estudio comparativo de técnicas para la segmentación automática de la materia blanca cerebral, proponiendo diferentes arquitecturas de segmentación como la U-Net y variantes de la misma. Se observan unas métricas mejores para una U-Net residual que, para las variantes implementadas, teniendo para U-Net un Dice Score de 0.86, para BConvLSTM U-Net un Dice Score de 0.79, para U-Net residual un Dice Score de 0.89 y para U-Net doble un Dice Score de 0.84.

Palabras clave: Redes neuronales convolucionales, segmentación de materia blanca cerebral, aprendizaje automático, U-Net.

Abstract

White matter has an important role in the brain structure, and an injury or degradation in this part, can cause severe damage in all brain system, therefore, medical image segmentation is an important process when using support software for injuries detection, neurodegradation, or neurosurgery evaluation. In this research, a comparative analysis of white matter segmentation techniques will be made, proposing different segmentation architectures such as U-Net and U-Net variants. Better metrics are observed for a residual U-Net, having a Dice Score for simple U-net of 0.86, for BConvLSTM U-Net a Dice Score of 0.79, for residual U-Net a Dice Score of 0.89, and for Double U-Net a Dice Score of 0.84.

Keywords: Convolutional Neural Network, White Matter Segmentation, machine learning, U-Net.

1. Introducción

La materia blanca cerebral está formada por tractos o fibras que interconectan distintas regiones del cerebro. Sin embargo, esta red no es solo una masa de fibras entrecruzadas, ya que existen distintos tractos que conectan las partes del cerebro, permitiendo a la información viajar de

una región a otra. Por las regiones que conectan, estos tractos pueden clasificarse en dos tipos principales: Fibras comisurales y fibras de asociación [1].

Como puede observarse en la figura 1, las fibras comisurales conectan regiones equivalentes entre los dos hemisferios del cerebro, por ejemplo: La comisura anterior, conecta las regiones olfatorias izquierda y derecha, mientras que la comisura posterior, conecta las fibras de los núcleos visuales del tronco encefálico. En general, estas fibras son muy pequeñas y pueden ser difíciles de encontrar, sin embargo, existe un grupo de fibras comisurales, que es por mucho el más grande: Corpus callosum, el cual se encuentra en el centro del cerebro. Por otro lado, las fibras de asociación, conectan áreas corticales en el mismo hemisferio. A su vez, estas se dividen en fibras de asociación largas, las cuales pueden presentarse de forma vertical u horizontal, y son conocidas como fascículos; y fibras de asociación cortas que tienen forma de U, las cuales conectan giros cerebrales adyacentes [1].

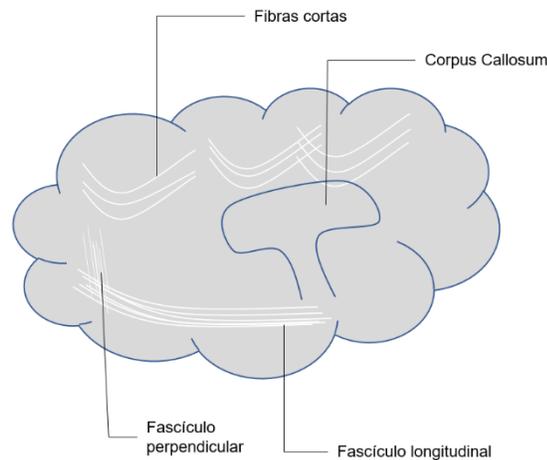


Figura 1. Clasificación general de la materia blanca cerebral.

Detectar a tiempo estas lesiones puede ser la salvación del paciente, y es por eso que los especialistas utilizan imágenes tomadas al cerebro como una herramienta visual para localizar daños o fallas en la materia blanca. Una forma de hacer esto con ayuda de un algoritmo computacional, es realizando una segmentación de cada uno de los cortes tomados del cerebro del sujeto, como puede apreciarse en la figura 2.

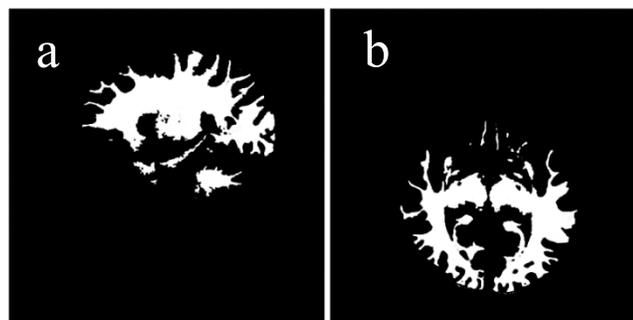


Figura 2. Segmentación de materia blanca cerebral, a) corte sagital y b) corte axial.



La segmentación de imágenes médicas, ha sido esencial para el diagnóstico asistido por computadora, ya que permite extraer de una imagen médica la región de interés (Region of interest ROI), y así poder realizar mediciones de manera más eficiente. La segmentación divide una imagen médica en distintas áreas, dependiendo de lo que se desea segmentar, por ejemplo, pueden segmentarse órganos completos, bordes, tumores, etc. Existe un gran interés en la segmentación automática de la materia blanca por la relación que existe entre la integridad de la estructura de los tractos y la neurodegeneración [2].

Los métodos basados en el procedimiento para visualizar los tractos neurales a través de resonancia magnética, mejor conocido como tractografía, son populares para realizar la segmentación de la materia blanca, sin embargo, dichos métodos no son óptimos para su utilización en la práctica clínica, ya que poseen una limitada consistencia, así como tiempos muy largos de procesamiento. Así mismo, existe la problemática debida a la posible presencia de múltiples enfermedades en cada sujeto. Muchos sujetos mayores de 60 años, presentan distintos grados de isquemia vascular cerebral, así como infartos crónicos, los cuales pueden manifestarse como lesiones en la materia blanca (white matter hyperintensities, WMH), en las imágenes por resonancia magnética [3].

En el año 2000, se hablaba ya de distintos métodos para realizar segmentaciones a imágenes médicas [4], métodos que pudieran englobarse en 8 diferentes categorías: umbralización [5], crecimiento de la región [6], clasificadores [7], técnicas de agrupamiento [8], modelos de campo aleatorio de Markov [9], redes neuronales artificiales [10], modelos deformables [11] y métodos guiados por atlas [12]. Una investigación del 2009 [13], propone el uso de algoritmos genéticos para la segmentación de imágenes médicas, aportando cierta flexibilidad en el procedimiento de segmentación automática. En 2010, otra investigación [14] desarrolla un modelo predictivo para realizar la extracción automática del contorno de objetos. Se han propuesto también modelos para realizar segmentaciones a imágenes en 3D, como los modelos estadísticos de forma (Statistical Shape Models SSM) [15].

Un artículo en el 2018 [16], publica sus resultados obtenidos con un método propuesto para mejorar la precisión de la segmentación de hiperintensidades de la materia blanca utilizando U-Net. El método propuesto es una técnica de post procesamiento, que aprovecha el descubrimiento del autor de que una combinación de umbrales y promediando las salidas de U-Nets con distintas inicializaciones aleatorias, pueden mejorar enormemente la precisión de hiperintensidades de la materia blanca cerebral. Ya que la técnica es independiente del modelo utilizado, el autor afirma que también puede ser utilizada en situaciones donde se hayan utilizado otros modelos de *Deep Learning* sobre todo cuando se adopta una inicialización aleatoria y el pre entrenamiento no sea posible. Los mejores resultados obtenidos por esta investigación fueron los siguientes: U-Net1 con un 0.8118 de DSC score, U-Net2 con un 0.8269 de DSC score, U-Net3 con un 0.7742 de DSC score, y aplicando el método del autor calculando el promedio de las máscaras binarias se obtuvo una imagen con un 0.8228 DSC score.

El mismo año, se publica otra investigación [17] en la cual se propone una red neuronal convolucional propuesta anteriormente para segmentar lesiones cerebrales graves, adaptada para segmentar hiperintensidades en la materia blanca. Los datos utilizados para el estudio, fueron obtenidos del repositorio público *Alzheimers Disease Neuroimaging Initiative*. Reportan un DSC máximo de 0.7292.

En 2018, una investigación [18] aborda el problema de segmentación de tractogramas cerebrales, el cual tiene importantes aplicaciones en la planeación de neurocirugías y tractometría. Al existir una crítica común sobre los métodos no supervisados utilizados para la segmentación del tracto, la investigación se enfoca en un método supervisado de segmentación, el cual se alimenta de atlas anatómicos y de tractos segmentados de diferentes sujetos. Dicho método segmenta un tracto de interés en el tractograma de un nuevo sujeto utilizando información previa. Su método está diseñado desde la perspectiva del problema de asignación lineal (Linear Assignment Problem, LAP)



obteniendo de segmentaciones de distintas regiones de la materia blanca un promedio de AUC (*Area under the Curve*) de 0.83.

En 2018, se desarrolla para una investigación [19], una herramienta rápida y precisa para la segmentación de materia blanca utilizando imágenes por resonancia magnética. El autor afirma que su método es mucho más rápido que cualquier otro método existente hasta entonces, y provee un Dice Score de 0.84. Se utiliza una red neuronal convolucional codificadora-decodificadora que procesa los datos y genera mapas de probabilidad.

En 2019, se desarrolló un método [20] basado en redes neuronales convolucionales para segmentar directamente los tractos de la materia blanca, obteniendo en su puntaje más alto para el tracto corticoespinal derecho un Dice Score de 0.76.

El mismo año, también se desarrolló un algoritmo automatizado [21] capaz de distinguir dichas enfermedades para cada individuo desde sus imágenes por resonancia magnética. Para resolver el problema entrenaron una red neuronal (3D U-Net) utilizando una base de datos del International Brain Tumor Segmentation (BraTS) 2018 dataset. Se obtuvo apenas un Dice Score de 0.42 para hiperintensidad de la materia blanca.

En 2020, se publicó una investigación [22] en la cual se propone una Red Generativa Adversaria (GAN) para la segmentación de materia blanca cerebral. El autor menciona de los algoritmos relacionados con aprendizaje profundo, las Redes Generativas Adversarias han obtenido un desempeño excelente en tareas de segmentación de imágenes. El resultado obtenido en el Dice Score por esta investigación fue de 0.817.

En el 2021, se propone también el uso de una GAN para la segmentación de materia blanca [23] pero enfocado a hiperintensidades. En cuanto a los resultados obtenidos, para el conjunto de pacientes con lesiones severas, se obtuvo un Dice Score de 0.712; para el conjunto de pacientes con lesiones moderadas se obtuvo un Dice Score de 0.494; y para el conjunto de pacientes con lesiones leves se obtuvo un Dice Score de 0.129.

En el mismo año, se publica una investigación [24] en la cual se propone la segmentación de la anomalía difusa de la materia blanca en infantes prematuros. El modelo ResU-Net tridimensional (3D) segmentó obteniendo un Dice de $0,907 \pm 0,041$ (desviación estándar) y una precisión equilibrada de $96,0\% \pm 2,1$, superando a múltiples modelos de aprendizaje profundo de pares. El modelo 3D ResU-Net que se entrenó con el conjunto de 98 sujetos, se probó más en un conjunto de validación de 28 sujetos y logró un Dice de $0,877 \pm 0,059$ y una precisión equilibrada de $92,3\% \pm 3,9$.

En la presente investigación, se propone una red convolucional U-Net, así como variantes de la misma para realizar comparaciones de segmentaciones de materia blanca a imágenes MRI de pacientes sanos.

2. Metodología

La base de datos utilizada para el estudio comparativo, fue proporcionada por el Laboratorio Nacional de Imagenología por Resonancia Magnética (LANIREM), ubicado en el Instituto de Neurobiología de la UNAM campus Juriquilla. Cuenta con 20 volúmenes de sujetos sanos en formato DICOM, cada volumen cuenta con un promedio de 256 imágenes de 512×512 . Como se puede apreciar en la figura 3, las imágenes son de cortes tanto sagitales como axiales.

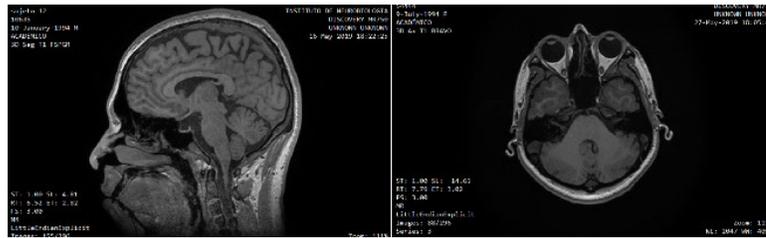


Figura 3. Imágenes DICOM del cerebro, en cortes sagital y axial de izquierda a derecha respectivamente.

Las imágenes proporcionadas por el laboratorio no se encontraban segmentadas, así que, para poder realizar apropiadamente los entrenamientos, se recurrió a un software especializado para la segmentación automática de la materia blanca a partir de las imágenes de la base de datos. El software utilizado para este propósito, es el Statistical Parametric Mapping, desarrollado por la University College of London UCL, el cual fue diseñado para el análisis de secuencias de imágenes cerebrales, y posee diversas herramientas para lograrlo, siendo una de ellas, la segmentación de materia blanca cerebral. Las máscaras obtenidas mediante SPM, fueron imágenes de 512x512 uint8 en formato png. El total de imágenes utilizadas en todos los experimentos, fue de 1184 imágenes cerebrales, con sus respectivas máscaras, haciendo que el modelo realizara una división de 80/20, 80% de las imágenes para entrenamiento, y 20% para validación de las métricas en cada época entrenada. El total de las 1184 imágenes, pertenecen a 4 sujetos, siendo dos conjuntos de imágenes en corte sagital y dos en corte axial.

Para esta investigación, se propusieron cuatro arquitecturas de entrenamiento: U-Net, BConvLSTM U-Net, U-Net Residual y U-Net Doble. U-Net [25] es una red neuronal convolucional, útil para realizar segmentaciones automáticas de imágenes utilizando bases de datos pequeñas, algo deseable cuando se trabaja con imágenes segmentadas por la dificultad de obtener bases de datos extensas. El diseño implementado para esta investigación, se puede apreciar en la figura 4.

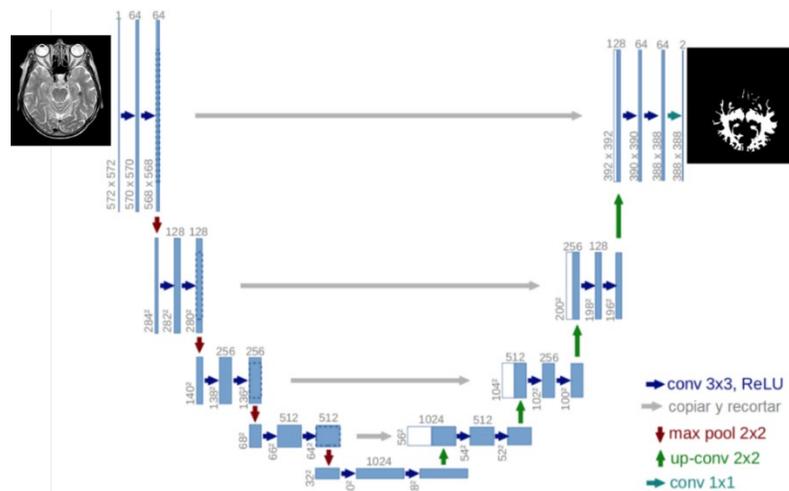


Figura 4. Arquitectura U-Net implementada en este proyecto.

Una variante U-Net, es la ya mencionada BConvLSTM U-Net [26], que puede observarse en la figura 5. Esta arquitectura realiza exactamente el mismo procedimiento que la arquitectura U-Net, con

la variante de utilizar LSTM [27](Long Short-Term Memory) bidireccional en las concatenaciones de cada nivel de la arquitectura. Esto supone una mayor precisión en la extracción de características por la metodología de LSTM. LSTM es una red neuronal recurrente, la cual tiene conexiones de retroalimentación, en donde es posible predecir la siguiente información basada en las conexiones anteriores. A diferencia de la mayoría de las redes neuronales recurrentes tradicionales, en donde hay ocasiones en las que se vuelve más complicado conectar con información que se encuentra en ciclos más alejados o “memoria más antigua” para obtener el contexto necesario para una predicción acertada, LSTM resuelve este problema siendo una red que permite conectar tanto con ciclos recientes como más antiguos [28].

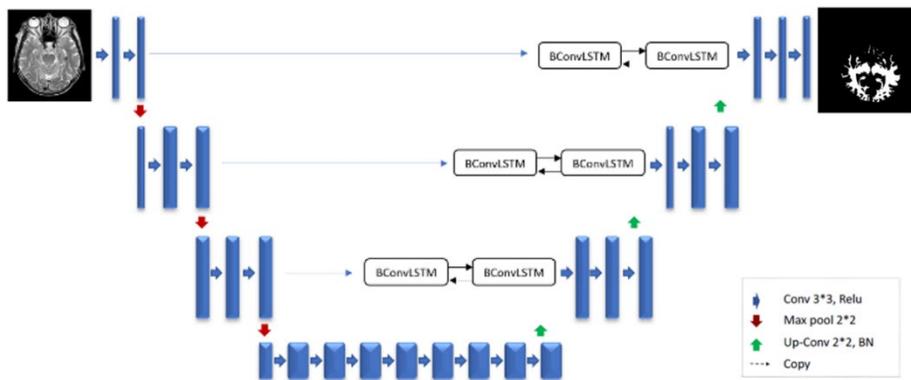


Figura 5. Diseño de la arquitectura U-Net con BConvLSTM implementada.

U-Net residual [29], utiliza la misma forma base de una U-Net simple, codificando y decodificando las imágenes para obtener características de entrenamiento, con la diferencia de que utiliza bloques residuales para intentar mejorar el flujo de la información en la arquitectura. En la figura 6 se puede observar el diseño implementado en la investigación, y en la figura 7 se observa con detalle la implementación de un bloque residual.

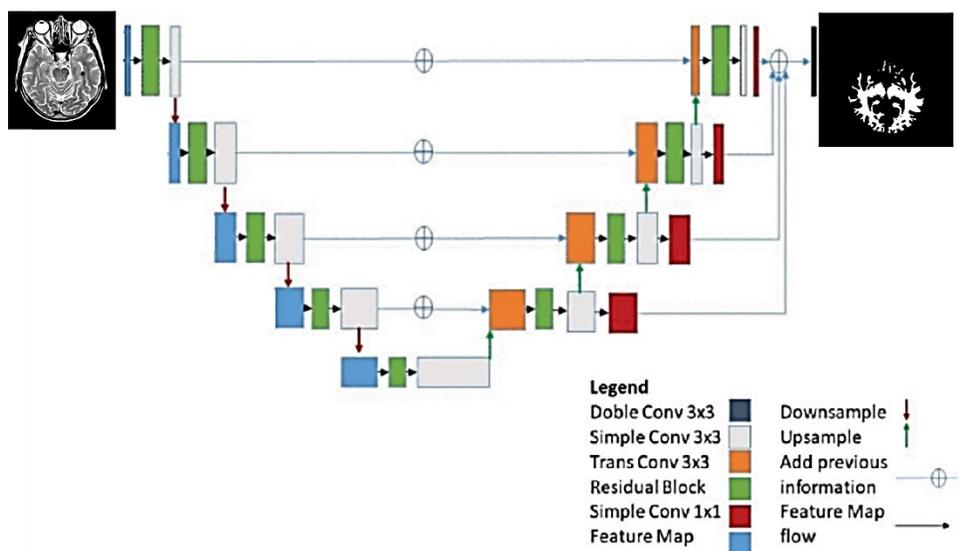


Figura 6. Diseño de la arquitectura U-Net residual implementada.

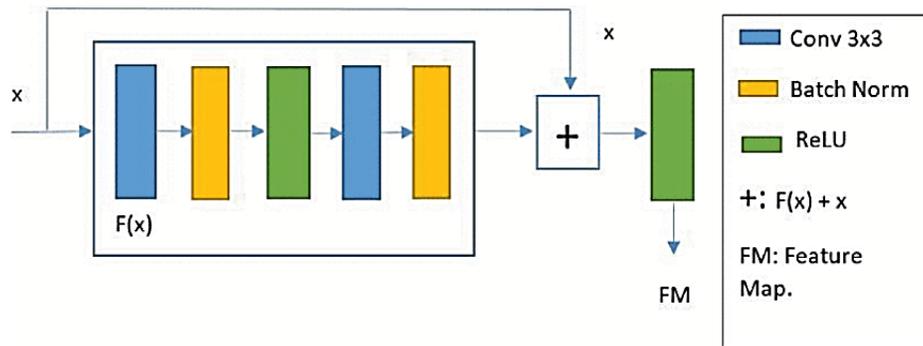


Figura 7. Diseño de un bloque residual utilizado para la implementación de U-Net residual.

La cuarta arquitectura utilizada es una U-Net doble [30] que hace uso de dos redes conectadas, una primer red con la forma de U-Net simple, y la salida de ésta primer red, se introduce a U-Net residual. Se puede observar en la figura 8 el diseño implementado en esta investigación.

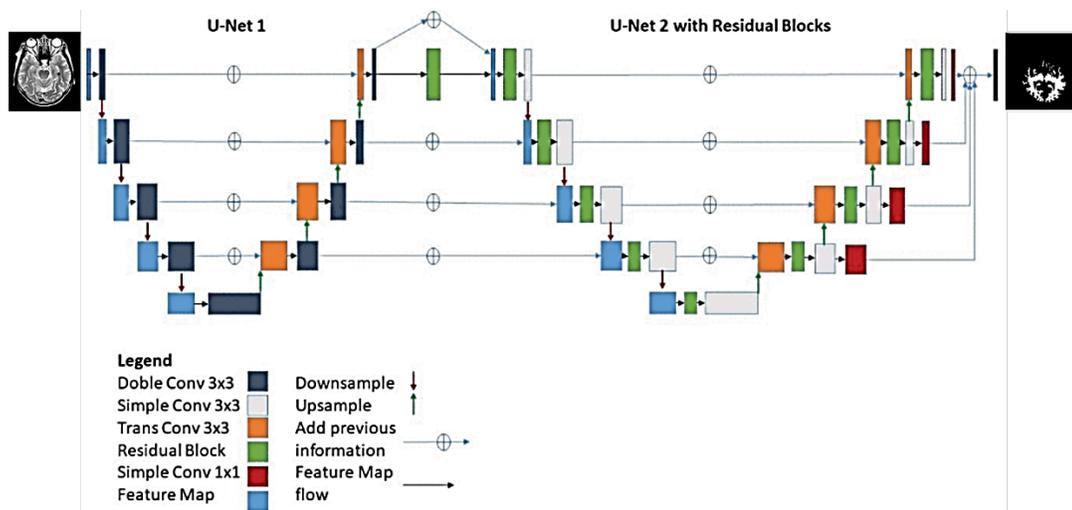


Figura 8. Diseño de la arquitectura U-Net doble implementada.

En cuanto a los hiperparámetros utilizados para las cuatro arquitecturas, pueden ser observados en la tabla 1. En ambas arquitecturas se utilizó la función de activación ReLU (1) [31] para los caminos de contracción y expansión, un optimizador adam [32] y la función de pérdida para una segmentación binarizada: binary crossentropy (2) [33]. Así mismo, para evitar un sobre entrenamiento, en ambas arquitecturas se utilizó una función para monitorear la pérdida generada en cada época, de tal forma que, si el valor de dicha pérdida no tiene diferencia en 3 épocas, el entrenamiento se detiene.

Función de activación ReLU. La función de activación ReLU activa una neurona solo cuando el valor de entrada (x) supera un umbral dado. Como se observa en la ecuación (1), si el umbral es 0, la función regresa el mismo valor cuando x es mayor o igual a cero, en caso de que sea negativo, regresa un 0, de este modo, tenemos una función de activación uno a uno. La activación ReLU mejora



la velocidad computacional para la etapa de entrenamiento y previene el problema de desvanecimiento del gradiente [30].

$$f(x) = \max(0, x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases} \quad (1)$$

$$H_p(\mathbf{q}) = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i)) \quad (2)$$

Tabla 1. Hiperparámetros utilizados en cada arquitectura.

Arquitectura	Función de activación	Optimizador	Función de pérdida
U-Net	ReLu ¹	adam	Binary crossentropy ²
BConvLSTM U-Net	ReLu ¹	adam	Binary crossentropy ²
U-Net Residual	ReLu ¹	adam	Binary crossentropy ²
U-Net Doble	ReLu ¹	adam	Binary crossentropy ²

El equipo utilizado para el entrenamiento de las arquitecturas, fue un equipo en la nube utilizando Google Colab, el cual cuenta con un GPU Tesla K80 con 2496 CUDA cores, 12GB GDDR5 VRAM, un CPU single core hyper threaded Xeon 2.3Ghz y 12GB de Ram.

El entorno de programación utilizado fue Python 3.7.6, haciendo uso de las bibliotecas TensorFlow 2.6.0 y Keras. 2.3.1.

La discusión más detallada de los resultados, puede observarse en la sección correspondiente. Cabe mencionar que las ejecuciones de todos los entrenamientos y utilización de las métricas para validar los modelos, se repitieron al menos 10 veces para cada arquitectura, y así comprobar que no existan variaciones importantes entre entrenamientos. Las métricas de calidad utilizadas para la evaluación de los resultados, se explican a continuación.

Para el cálculo de cada una de las métricas [34], es necesario etiquetar los conjuntos de datos predichos por la arquitectura, según su comparación contra los datos verdaderos utilizados para el entrenamiento. De esta forma quedan de la siguiente manera:

- Positivos que son correctamente predichos por el modelo como positivos, son verdaderos positivos (TP).
- Positivos que son incorrectamente predichos por el modelo como negativos, son falsos negativos (FN).
- Negativos que son correctamente predichos por el modelo como negativos, son verdaderos negativos (TN).
- Negativos que son incorrectamente predichos por el modelo como positivos, son falsos positivos (FP).

Accuracy. Muchos autores consideran que la métrica *accuracy* es el estándar desde donde partir, ya que es la medida más directa que puede calcularse. Se puede observar el cálculo de *accuracy* en la ecuación 3 [34]. Es una métrica bastante intuitiva y natural, ya que se calcula la



proporción de todas las predicciones correctas del modelo, sin embargo, no es apropiada en muchos casos. Una desventaja significativa, es cuando las variables destino no se encuentran balanceadas, lo cual sucede en la mayoría de las ocasiones.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (3)$$

Recall. *Recall* es la proporción de elementos identificados correctamente como positivos (TP) del total de positivos verdaderos. Se puede observar el cálculo en la ecuación 4 [34].

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

Precision. Es la proporción de elementos identificados correctamente como positivos (TP) con respecto a todos los positivos. Se puede observar el cálculo en la ecuación 5 [34].

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

F1 Score o Dice Similarity Coefficient. En ocasiones, no es suficiente con calcular solo el *precision* o el *recall*, ya que ambas métricas son igualmente importantes, en ese caso, es posible utilizar la métrica *F1 Score* o *DSC Dice Similarity Coefficient* para obtener la media armónica entre estos dos valores. Esta métrica es la más utilizada para evaluar segmentación de imágenes médicas [17]. Se puede observar el cálculo en la ecuación 6 [34].

$$F1Score/DSC = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (6)$$

3. Resultados

Las métricas de calidad que se obtuvieron al finalizar cada uno de los entrenamientos, se muestran en las tablas 2 y 3. En la tabla 2, se muestran los resultados con las imágenes utilizadas para el entrenamiento, mientras que en la tabla 3 se muestran los resultados al evaluar el conjunto de imágenes utilizadas por el modelo para la validación.

Tabla 2. Métricas de calidad obtenidas en cada una de las arquitecturas para las imágenes de entrenamiento.

Arquitectura	Accuracy	Recall	DSC	Épocas	Tiempo de entrenamiento
U-Net	0.98	0.92	0.93	15	9 minutos 49 segundos
BConvLSTM U-Net	0.98	0.88	0.88	12	109 minutos 12 segundos
Residual U-Net	0.99	0.96	0.96	18	10 minutos 22 segundos
Double U-Net	0.98	0.91	0.91	15	22 minutos 25 segundos

Tabla 3. Métricas de calidad obtenidas en cada una de las arquitecturas para las imágenes de validación.

Arquitectura	Accuracy	Recall	DSC
U-Net	0.98	0.91	0.86
BConvLSTM U-Net	0.99	0.81	0.79
Residual U-Net	0.98	1	0.89
Double U-Net	0.98	0.99	0.84

Las figuras 9 a 13, muestran una comparativa de las métricas obtenidas en cada época de la arquitecturas utilizadas: accuracy, recall y Dice Similarity Coefficient respectivamente, tanto para las imágenes de entrenamiento como para las imágenes de validación.

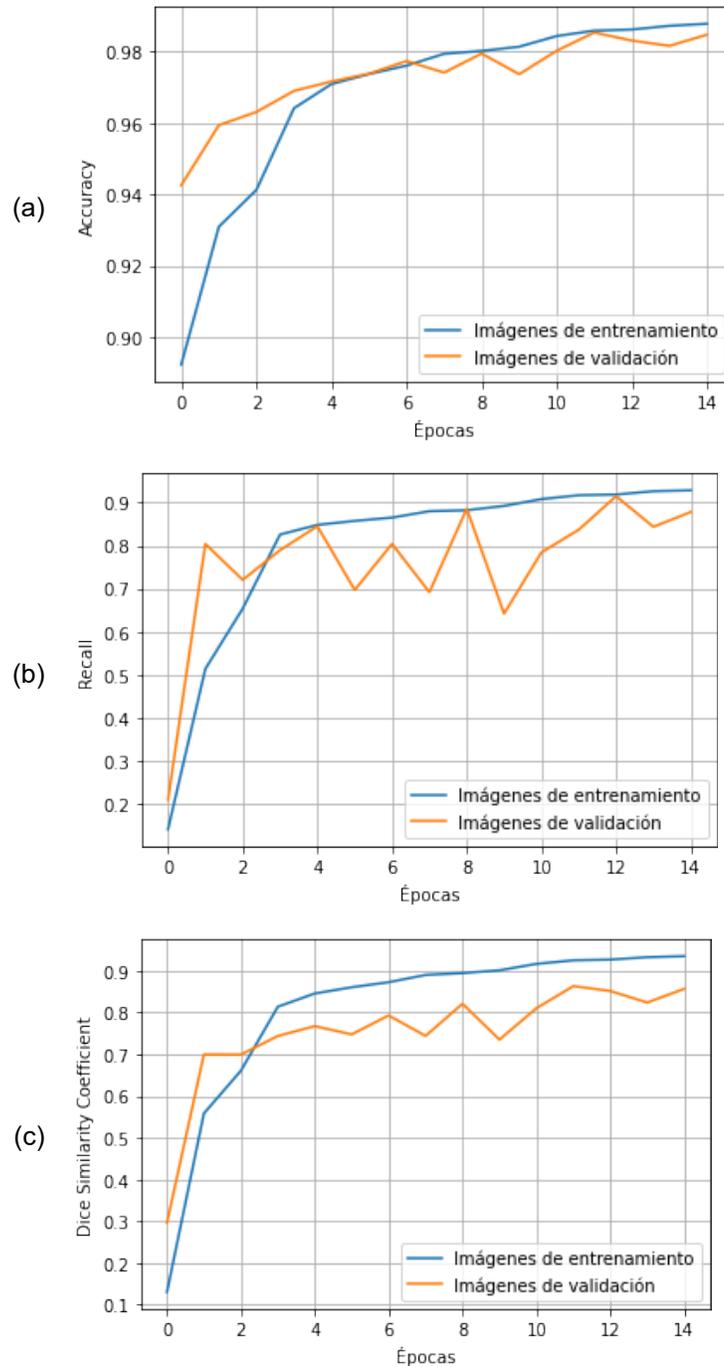


Figura 9. (a) Accuracy obtenido en cada época de U-Net para imágenes de entrenamiento e imágenes de validación.; (b) Recall obtenido en cada época de U-Net para imágenes de entrenamiento e imágenes de validación.; (c) Dice Similarity Coefficient obtenido en cada época de U-Net para imágenes de entrenamiento e imágenes de validación.

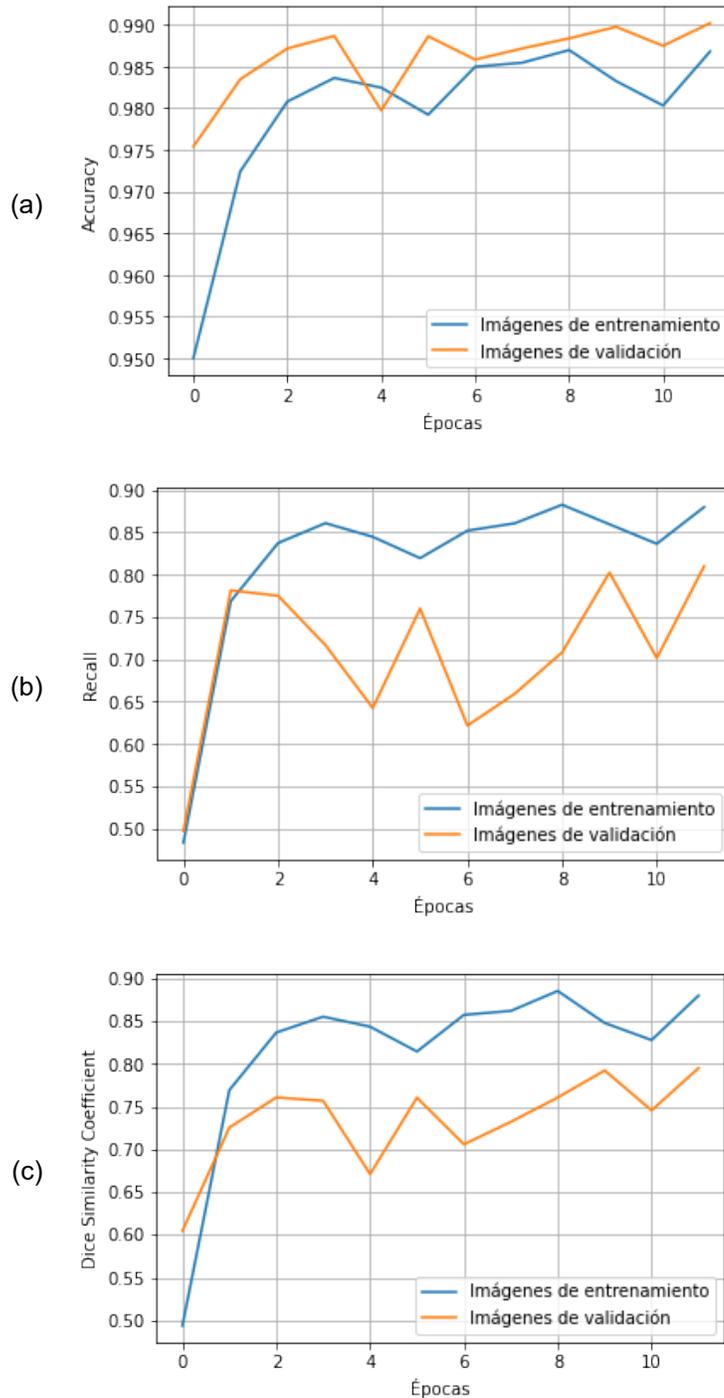


Figura 10. (a) Accuracy obtenido en cada época de BConvLSTM U-Net para imágenes de entrenamiento e imágenes de validación.; (b) Recall obtenido en cada época de BConvLSTM U-Net para imágenes de entrenamiento e imágenes de validación.; (c) Dice Similarity Coefficient obtenido en cada época de BConvLSTM U-Net para imágenes de entrenamiento e imágenes de validación.

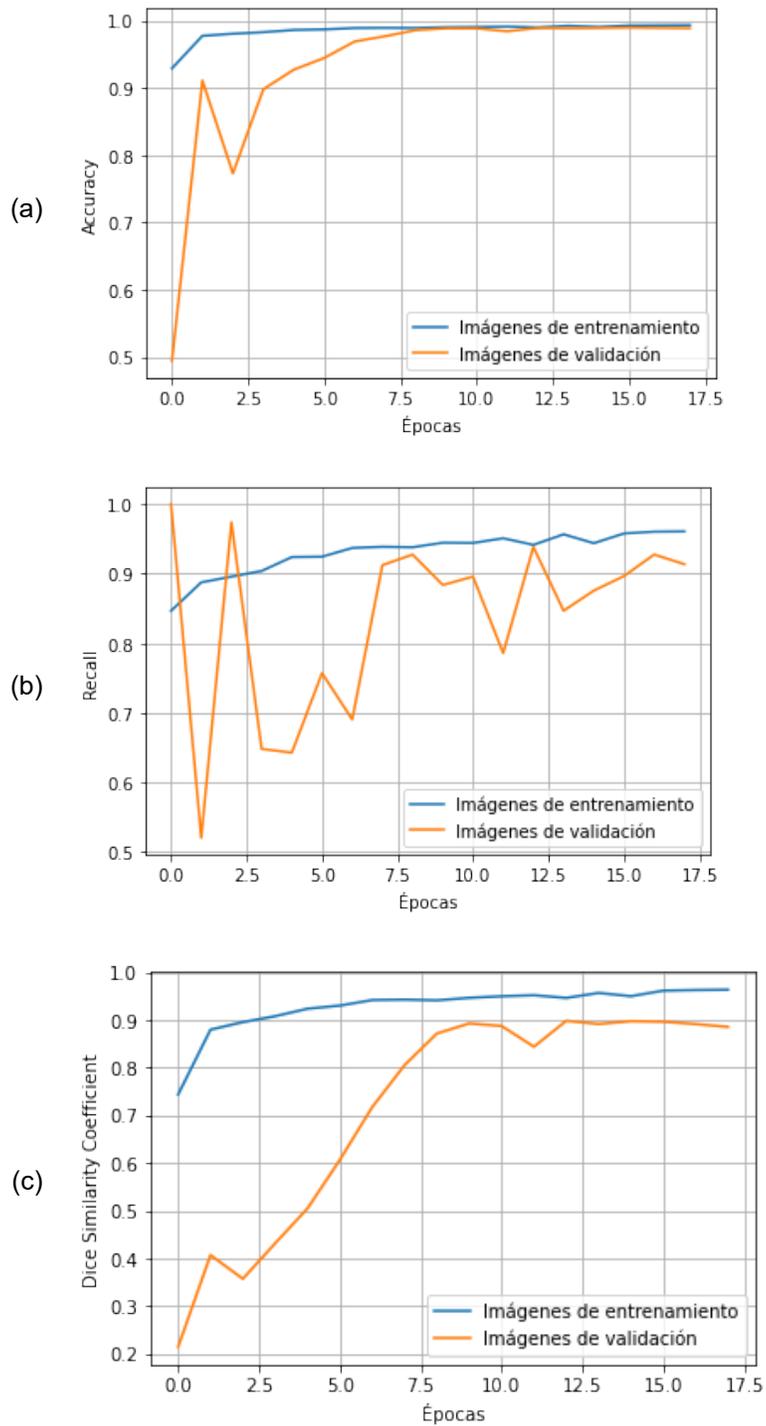


Figura 11. (a) Accuracy obtenido en cada época de U-Net Residual para imágenes de entrenamiento e imágenes de validación.; (b) Recall obtenido en cada época de U-Net Residual para imágenes de entrenamiento e imágenes de validación.; (c) Dice Similarity Coefficient obtenido en cada época de U-Net Residual para imágenes de entrenamiento e imágenes de validación.

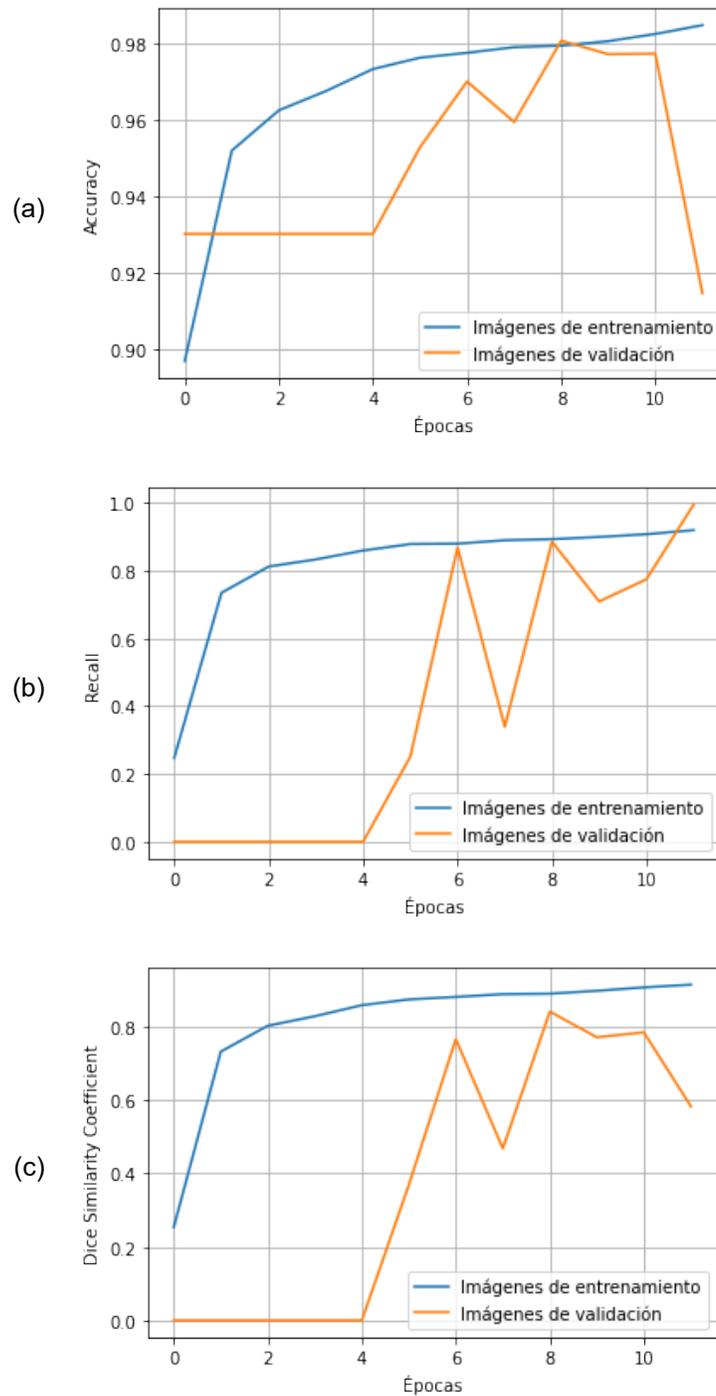


Figura 11. (a) Accuracy obtenido en cada época de U-Net Doble para imágenes de entrenamiento e imágenes de validación.; (b) Recall obtenido en cada época de U-Net Doble para imágenes de entrenamiento e imágenes de validación.; (c) Dice Similarity Coefficient obtenido en cada época de U-Net Doble para imágenes de entrenamiento e imágenes de validación.

Las figuras 12 a 19 muestran una comparativa de las segmentaciones obtenidas para una imagen de prueba tanto de corte sagital como axial, por las distintas arquitecturas utilizadas en esta

investigación, en donde se observan: imagen original, el *ground truth*, y la predicción del modelo entrenado.

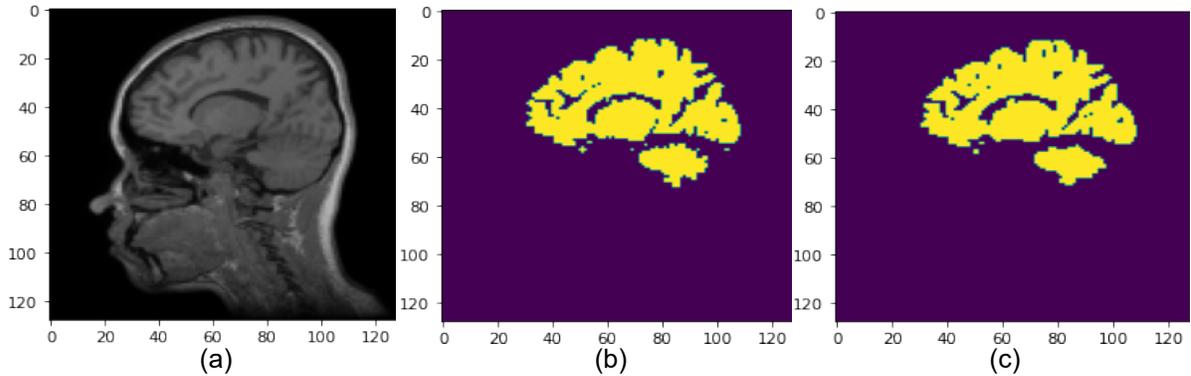


Figura 12. (a) Imagen de corte sagital de un cerebro humano.; (b) *ground truth* utilizado para el entrenamiento de la arquitectura U-Net.; (c) Predicción realizada por la arquitectura U-Net .

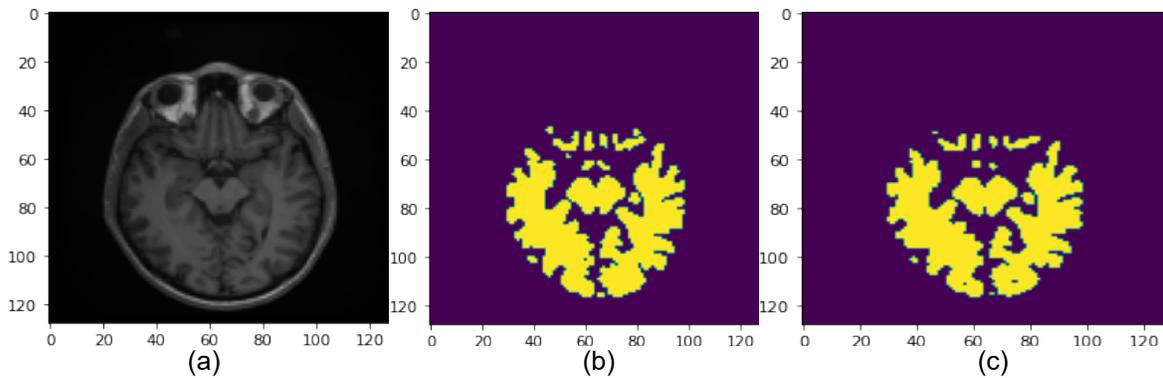


Figura 13. (a) Imagen de corte axial de un cerebro humano.; (b) *ground truth* utilizado para el entrenamiento de la arquitectura U-Net.; (c) Predicción realizada por la arquitectura U-Net .

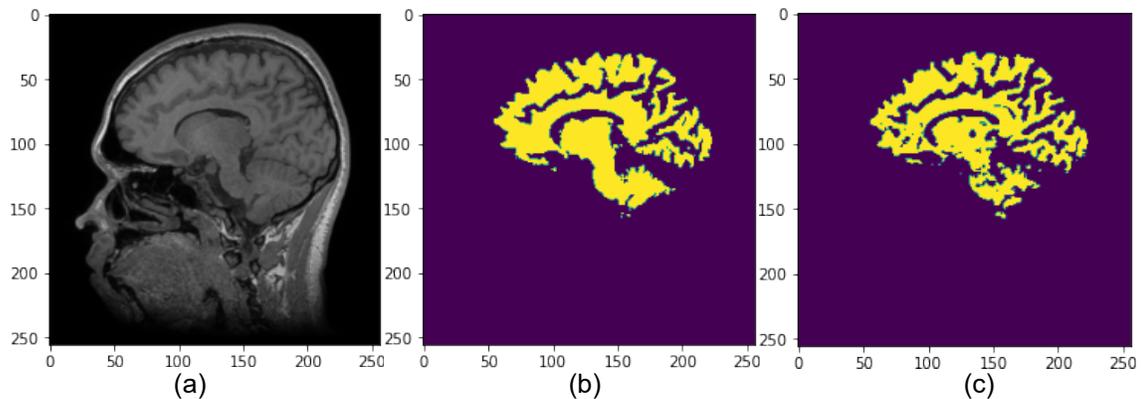


Figura 14 a) Imagen de corte sagital de un cerebro humano.; (b) *ground truth* utilizado para el entrenamiento de la arquitectura BConvLSTM U-Net.; (c) Predicción realizada por la arquitectura BConvLSTM U-Net .

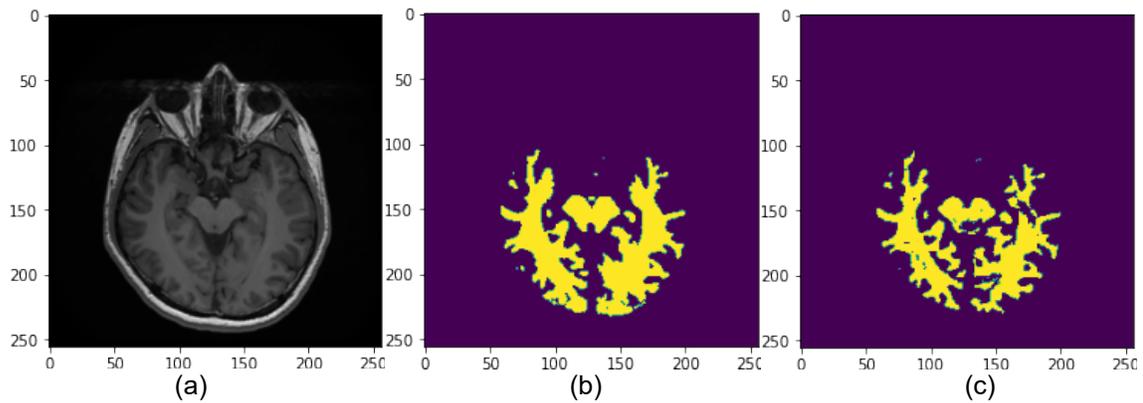


Figura 15. (a) Imagen de corte axial de un cerebro humano.; (b) *ground truth* utilizado para el entrenamiento de la arquitectura BConvLSTM U-Net.; (c) Predicción realizada por la arquitectura BConvLSTM U-Net .

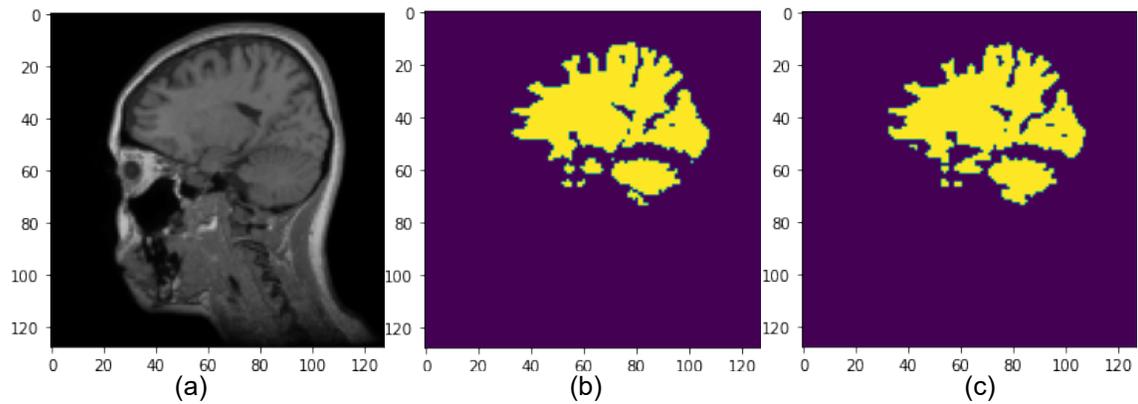


Figura 16 (a) Imagen de corte sagital de un cerebro humano.; (b) *ground truth* utilizado para el entrenamiento de la arquitectura Residual U-Net.; (c) Predicción realizada por la arquitectura Residual U-Net.

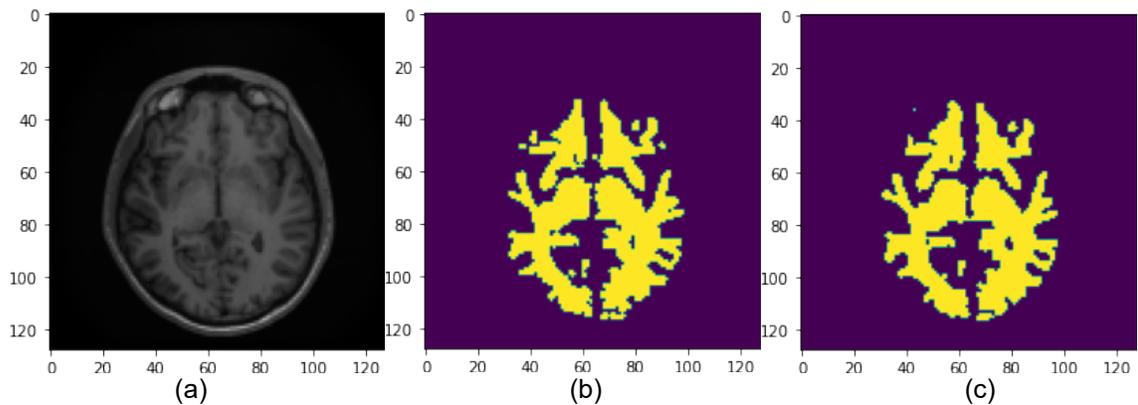


Figura 17. (a) Imagen de corte axial de un cerebro humano.; (b) *ground truth* utilizado para el entrenamiento de la arquitectura Residual U-Net.; (c) Predicción realizada por la arquitectura Residual U-Net.

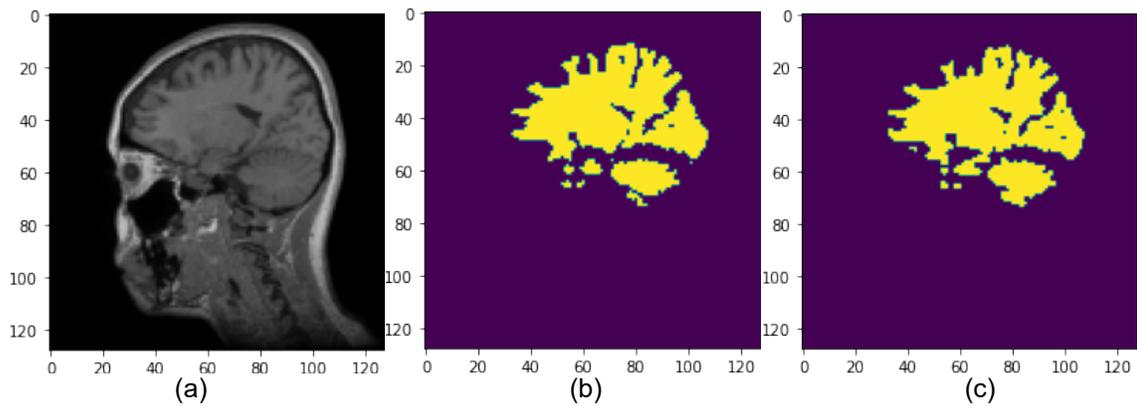


Figura 18. (a) Imagen de corte sagital de un cerebro humano.; (b) *ground truth* utilizado para el entrenamiento de la arquitectura Double U-Net.; (c) Predicción realizada por la arquitectura Double U-Net .

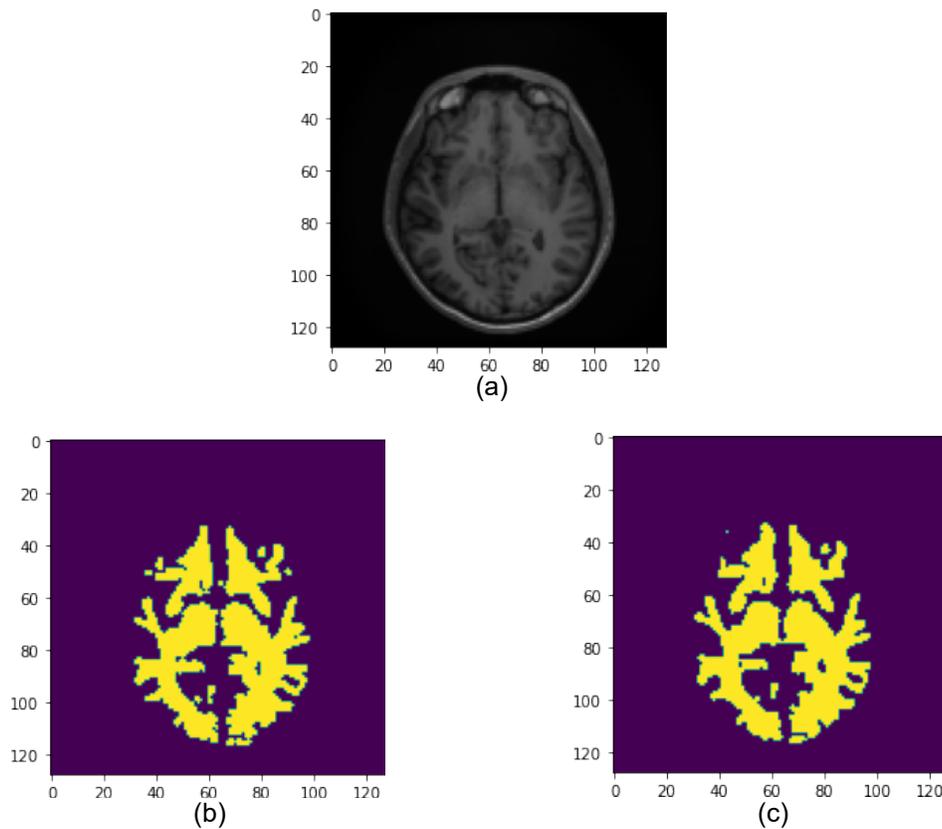


Figura 19. (a) Imagen de corte axial de un cerebro humano.; (b) *ground truth* utilizado para el entrenamiento de la arquitectura Double U-Net.; (c) Predicción realizada por la arquitectura Double U-Net .

Uno de los objetivos de esta investigación era realizar una comparativa entre las arquitecturas implementadas para la segmentación de la materia blanca, sin embargo, es importante realizar también una comparativa con respecto al estado del arte investigado.



En esta sección se muestra la tabla 4 en la que se pueden apreciar los resultados obtenidos por investigaciones anteriores, así como los obtenidos por esta investigación.

Hay que tener en consideración, que, aunque *Dice Similarity Coefficient* (DSC) es una de las métricas más utilizadas para evaluar la segmentación de imágenes médicas, algunos autores presentan en sus resultados otro tipo de métricas dependiendo del método implementado en sus investigaciones.

Tabla 4. Resultados y métodos utilizados en investigaciones anteriores e investigación del presente trabajo.

Autor	Resultados
F. Admiraal- Behloul, 2005	ICC de 0.99. La precisión de la segmentación se evaluó en 100 pacientes obteniendo un ICC de 0.98 y un índice de similitud de 0.75.
Y. Zhang, 2018	U-Net1 = 0.8118 DSC, U-Net2 = 0.8269 DSC, U-Net3 = 0.7742 DSC, y aplicando el método del autor calculando el promedio de las máscaras binarias se obtuvo una imagen con un 0.8228 DSC.
M. F. Rachmadi, 2018	0.7292 DSC.
N.Sharmin, 2018	NNDRMAM = 0.81 AUC , LAP = 0.90 AUC. Resultados más altos obtenidos, siendo estos para el tracto del ventrículo izquierdo y un promedio de 0.83 AUC usando LAP.
J. Wasserthal, 2018	DSC: 0.84.
B. Li, 2019	Corpus Callosum DCS = 0.67; Mean K = 0.79; Tracto corticoespinal derecho, DSC = 0.76; Mean K = 0.84.
D. Rudie, 2019	DSC = 0.92 para el tumor completo
O. Kyeong, 2020	DSC = 0.817
O. Kyeong, 2021	Lesiones severas: DSC = 0.712, Lesiones moderadas: DSC = 0.492, Lesiones leves: DSC = 0.129.
H. Li, 2021	DSC = 0,877 ± 0,059 y una precisión equilibrada de 92,3% ± 3.9. con 28 imágenes de validación.
Propuesta	U-Net, DSC = 0.86. BConvLSTM U-Net, DSC = 0.79. U-Net Residual, DSC = 0.89. U-Net Doble, DSC = 0.84. Resultados obtenidos con imágenes de validación.



4. Conclusiones

En este artículo, se presentaron diversas arquitecturas para realizar un entrenamiento de segmentación automática de la materia blanca cerebral, utilizando una base de datos de imágenes reales de sujetos sanos. El preprocesamiento de estas imágenes, fue en este caso, la etapa que mas tiempo requirió, ya que conseguir imágenes médicas es un proceso delicado y no tan sencillo. Posteriormente a la obtención de la base de datos, una segmentación de las imágenes con software especializado, fue necesaria para obtener arquitecturas.

Al observar los resultados de los entrenamientos, se puede notar que, en este caso, una arquitectura U-Net residual obtiene métricas de calidad mejores con respecto a las otras arquitecturas implementadas.

Respecto a estado del arte, se observa que se obtienen resultados muy similares, sin embargo, mas que hacer una comparativa con otras investigaciones, uno de los hallazgos mas relevantes en esta investigación, fue al momento de comparar las distintas arquitecturas utilizadas para el entrenamiento. Encontramos, que al menos en este caso particular, una arquitectura U-Net residual obtiene mejores resultados en segmentación de materia blanca que las variantes de U-Net implementadas. Pudiera discutirse mas a fondo el motivo de éstos resultados, realizando otro tipo de experimentos para encontrar de manera puntual la razón de este comportamiento y reafirmar si, para la mayoría de entrenamientos similares al realizado en esta investigación, sea la mejor opción una red U-Net residual.

Referencias

- [1] Mancall E., Brock D., "Section V / The cerebrum in Gray's Clinical Neuroanatomy, The Anatomic Basis for Clinical Neuroscience", *Elsevier Saunders*, USA, 1st ed., 2011.
- [2] Yanhui G., Amira S., "Neutrosphic sets in dermoscopic medical image segmentation", *Neutrosophic Set in Medical Image Analysis*, Academic Press , pp. 229-243, 2019.
- [3] Zhou Z., Rahman M., Tajbakhsh N., Liang J. (2018) "UNet++: A Nested U-Net Architecture for Medical Image Segmentation", *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Lecture Notes in Computer Science, vol 11045, 2018.
- [4] Dzung L., Chenyang X., Prince J., "Current Methods in Medical Image Segmentation", *Annual Review of Biomedical Engineering*, vol. 2, pp. 322, 2000.
- [5] Mehmet S., Bülent S., "Survey over image thresholding techniques and quantitative performance evaluation," *J. Electron* , 2004.
- [6] Pohle R., Toennies K., "Segmentation of medical images using adaptive region growing," *Medical Imaging 2001: Image Processing*, 2001.
- [7] Rohlfing T., Maurer C, "Multi-classifier framework for atlas-based image segmentation", *Pattern Recognition Letters*, Volume 26, Issue 13, pp. 2070-2079, 2005.
- [8] Coleman, G. B., & Andrews, H. C., "Image segmentation by clustering", *Proceedings of the IEEE*, 67(5), pp. 773-785, 1979.
- [9] Chen, X., Zheng, C., Yao, H., Wang, B., "Image segmentation using a unified Markov random field model.", *IET Image Processing*, 11(10), pp. 860-869, 2017.
- [10] Suzuki K., "Medical image segmentation using artificial neural networks," *Artificial Neural Networks - methodological advances and biomedical applications*, Rijeka: InTech, pp. 121–130, 2011.
- [11] Xu C., Pham D., Prince J., "Image segmentation using deformable models," *Handbook of Medical Imaging*, Volume 2. Medical Image Processing and Analysis, pp. 129–174, 2000.
- [12] Arabi H., Koutsouvelis N., Rouzard M., Miralbell T., Zaidi H., "Atlas-guided generation of pseudo-ct images for MRI-only and hybrid PET–MRI-guided radiotherapy treatment planning," *Physics in Medicine and Biology*, vol. 61, no. 17, pp. 6531–6552, 2016.



- [13] Maulik U., "Medical Image Segmentation Using Genetic Algorithms," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 2, pp. 166-173, 2009.
- [14] Shen T., Li H. and Huang X., "Active Volume Models for Medical Image Segmentation," *IEEE Transactions on Medical Imaging*, vol. 30, no. 3, pp. 774-791, 2011.
- [15] Heimann T., Meinzer H., "Statistical shape models for 3D medical image segmentation: A review", *Medical Image Analysis*, Volume 13, Issue 4, Pages 543-563, 2009.
- [16] Y. Zhang, W. Chen, Y. Chen, and X. Tang, "A post-processing method to improve the white matter hyperintensity segmentation accuracy for randomly-initialized u-net," in 2018 IEEE 23rd International Conference on Digital Signal Processing (DSP), pp. 1–5, IEEE, 2018.
- [17] M. F. Rachmadi, M. d. C. Valdes-Hernandez, M. L. F. Agan, C. Di Perri, T. Komura, A. D. N. Initiative, et al., "Segmentation of white matter hyperintensities using convolutional neural networks with global spatial information in routine clinical brain mri with none or mild vascular pathology," *Computerized Medical Imaging and Graphics*, vol. 66, pp. 28–43, 2018.
- [18] Sharmin N., Olivetti E., Avesani P., "White Matter Tract Segmentation as Multiple Linear Assignment Problems", *Frontiers in Neuroscience*, vol. 11, pp. 754, 2018.
- [19] Wassserthal J., Neher P., Maier-Hein K., "TractSeg - Fast and accurate white matter tract segmentation", *NeuroImage*, vol. 183, pp. 239-253, 2018.
- [20] Li B., de Groot M., Vernooij M., Ikram M., Niessen W., Bron E. "Reproducible White Matter Tract Segmentation Using 3D U-Net on a Large-scale DTI Dataset," *Machine Learning in Medical Imaging Lecture Notes in Computer Science*, vol. 11046, pp 205-213, 2019.
- [21] Rudie J., Weiss D., Saluja R., Rauschecker A., Wang J., Sugrue L., Bakas S., Colby J., "Multi-Disease Segmentation of Gliomas and White Matter Hyperintensities in the BraTS Data Using a 3D Convolutional Neural Network", *Frontiers in Computational Neuroscience*, vol. 13, pp. 84, 2019.
- [22] K. T. Oh, S. Lee, H. Lee, M. Yun, and S. K. Yoo, "Semantic segmentation of white matter in fdg-pet using generative adversarial network," *Journal of Digital Imaging*, vol. 33, no. 4, p. 816–825, 2020.
- [23] K. T. Oh, D. Kim, B. S. Ye, S. Lee, M. Yun, and S. K. Yoo, "Segmentation of white matter hyperintensities on 18f-fdg pet/ct images with a generative adversarial network," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 48, no. 11, p. 3422–3431, 2021.
- [24] H. Li, M. Chen, J. Wang, V. S. P. Illapani, N. A. Parikh, and L. He, "Automatic segmentation of diffuse white matter abnormality on t2-weighted brain mr images using deep learning in very preterm infants," *Radiology: Artificial Intelligence*, vol. 3, no. 3, p. e200166, 2021.
- [25] Ronneberger O., Fischer P., Brox T., "U-net: Convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9351, pp. 234–241, 2015.
- [26] Azad R., Asadi-Aghbolaghi M., Fathy M., Escalera S., "Bi-Directional ConvLSTM U-Net with Densley Connected Convolutions", *Electrical Engineering and Systems Science*, eprint. 1909.00166, 2019.
- [27] Hochreiter S., Schmidhuber J., "Long short-term memory", *Neural Computation*. Vol. 9, pp. 1735–1780, 1997.
- [28] Fukushima, K.; Miyake, S., "Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition", *Competition and cooperation in neural nets*, Springer: 267–285. 1982
- [29] Khanna A., Narendra D., Gupta S., Semwal A., "A deep Residual U-Net convolutional neural network for automated lung segmentation", *Computed Tomography Images, Biocybernetics and Biomedical Engineering*, vol. 40, pp. 1314-1327, 2020.
- [30] Francia A., Pedraza C., Aceves M, Tovar S., "Chaining a U-Net With a Residual U-Net for Retinal Blood Vessels Segmentation," *IEEE Access*, vol. 8, pp. 38493-38500, 2020.
- [31] Banerjee C., Mukherjee T., Pasiliao E., "An empirical study on generalizations of the ReLU activation function," *Proceedings of the 2019 ACM Southeast Conference*, 2019.
- [32] Diederik P., Ba J., "Adam: A Method for Stochastic Optimization", *arXiv*, 2017.
- [33] Mannor S, Peleg D, Rubinstein R., "The cross entropy method for classification.", *Proceedings of the 22nd international conference on Machine learning*, pp. 561-568, 2005.



- [34] D. Chicco and G. Jurman, “The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation,” *BMC genomics*, vol. 21, no. 1, pp. 1–13, 2020.

Autores

Eduardo Daniel Posadas Gamez. Ingeniero en Software por la Universidad Autónoma de Querétaro, con experiencia de 3 años en desarrollo de software. Actualmente estudiante de la Maestría en Ciencias en Inteligencia Artificial en la Facultad de Ingeniería de la Universidad Autónoma de Querétaro. Su investigación trata sobre la segmentación de la materia blanca cerebral por medio de aprendizaje automático.

Saúl Tovar Arriaga. Obtuvo su grado de Licenciatura en Ingeniería en Electrónica en el Instituto Tecnológico de Querétaro, su Maestría en Ciencias en Mecatrónica en la Universidad de Siegen, Alemania, y su Doctorado en Ciencias Biomédicas en la Universidad de Erlangen-Nuremberg, Alemania. Actualmente es profesor de tiempo completo e investigador en la Universidad Autónoma de Querétaro. Sus intereses de investigación incluyen robótica médica, diagnóstico automatizado por imagen y visión por computadora.

Rafael Ortiz Peregrino. Ingeniero en Mecatrónica por la Universidad Politécnica de Querétaro, Maestro en Ciencias en Inteligencia Artificial por la Universidad Autónoma de Querétaro. Se dedica a la investigación en la Universidad Autónoma de Querétaro.

Erick Humberto Pasaye Alcaraz. Técnico Académico Titular «A» T. C., Unidad de Resonancia Magnética, Instituto de Neurobiología UNAM, Campus Juriquilla. Licenciado en Ciencias Físico-Matemáticas por Universidad Michoacana de San Nicolás, Maestro en Ciencias Físicas por Instituto de Física y Doctor en Ciencias Biomédicas por el Instituto de Neurobiología. El doctor Pasaye estuvo 5 años como físico del área de resonancia magnética de Instituto Nacional de Neurología y Neurocirugía, 7 años como académico titular «A» de la unidad de Resonancia Magnética del Instituto de Neurobiología, UNAM. Ha escrito 1 libro y 2 capítulos de libro, ha hecho publicaciones internacionales indizadas con arbitraje, publicaciones nacionales arbitradas, Congresos nacionales e internacionales y ha impartido 33 cursos como docente.

Agradecimientos

Por su apoyo a la realización de esta investigación, se agradece su contribución a los siguientes:

Por parte del Laboratorio Nacional de Visualización Científica Avanzada, a Luis A. Aguilar, Jair S. García y Alejandro De León. Por parte de la Unidad de Resonancia Magnética del INB UNAM Juriquilla, al Dr. Erick Pasaye y al Dr. Cesar A. Dominguez.