

IMPLEMENTACIÓN DE UN ALGORITMO DE SEGMENTACIÓN DE ACCIDENTES CEREBROVASCULARES EN IMÁGENES DE RESONANCIA MAGNÉTICA UTILIZANDO REDES CONVOLUCIONALES

*IMPLEMENTATION OF A STROKE SEGMENTATION
ALGORITHM IN MAGNETIC RESONANCE IMAGING
USING CONVOLUTIONAL NETWORKS*

Oscar Cervantes Chavarría

Instituto Politécnico Nacional, México
ocervantesc2200@alumno.ipn.mx

José Luis Díaz Reséndiz

Instituto Politécnico Nacional, México
jdiazr2100@alumno.ipn.mx

Volodymyr Ponomaryov

Instituto Politécnico Nacional, México
vponomar@ipn.mx

Recepción: 26/septiembre/2023

Aceptación: 24/abril/2024

Resumen

Un accidente cerebrovascular es una emergencia médica que puede provocar complicaciones graves si no se detecta a tiempo. De modo que, si se diagnóstica con anticipación aumenta la posibilidad de que el paciente pueda sobrevivir. La detección de las lesiones por accidente cerebrovascular isquémico mediante imágenes de resonancia magnética (RM) es fundamental para un diagnóstico preciso y decisivo.

En este artículo, se propone desarrollar un algoritmo mediante técnicas de procesamiento de imágenes y una arquitectura basada en U-Net para realizar la segmentación automática de accidentes cerebrovasculares a partir de IRM que permitan identificar la zona de la lesión. Se utilizó la base de datos ISLES 2015, dos optimizadores (SGD y AdamW) y tres funciones de pérdida (Dice, Focal, GDFL) para evaluación del método propuesto obteniendo mejores resultados con AdamW

y GDFL con un valor de intersección sobre unión de 0.791 que compite con métodos del estado del arte.

Palabras Clave: Accidente cerebrovascular, Generalized dice loss, Imágenes multimodales, Segmentación de imágenes de resonancia magnética, U-Net.

Abstract

A stroke is a medical emergency that can lead to serious complications if not detected early. Thus, early diagnosis increases the patient's chance of survival. Detection of ischemic stroke lesions by magnetic resonance imaging (MRI) is essential for accurate and decisive diagnosis.

In this article, we propose to develop an algorithm using image processing techniques and a U-Net based architecture to perform stroke segmentation from MRI to identify the lesion area. The ISLES 2015 database, two optimizers (SGD and AdamW) and three loss functions (Dice, Focal, GDFL) were used to evaluate the performance of the proposed method obtaining better results with AdamW and GDFL with an intersection over union value of 0.791 that compete with State-of-the-art methods.

Keywords: *Generalized dice loss, Magnetic resonance image segmentation, Multimodal imaging, Stroke, U-Net.*

1. Introducción

Una enfermedad vascular cerebral (EVC) es una alteración en las neuronas, que provoca la disminución del flujo sanguíneo en el cerebro, acompañada de alteraciones cerebrales de manera momentánea o permanente [IMSS, 2015].

Los accidentes se clasifican en dos tipos: isquémica y hemorrágica (Figura 1):

- **Hemorrágica:** Se presenta cuando el vaso sanguíneo se rompe, lo que causa un sangrado dentro del cerebro.
- **Isquémica:**
 - ✓ Isquémica cerebral transitoria: Se produce por la interrupción del flujo sanguíneo, los síntomas duran pocos minutos, por lo que el paciente se recupera totalmente en menos de 24 horas.

- ✓ **Infarto cerebral:** Se produce cuando la interrupción del flujo sanguíneo es causada por una obstrucción, que puede ser causada por grasa o sangre, en un vaso sanguíneo.

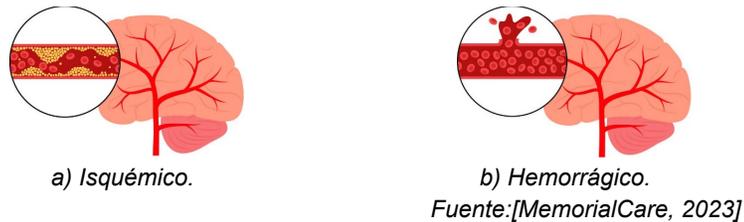
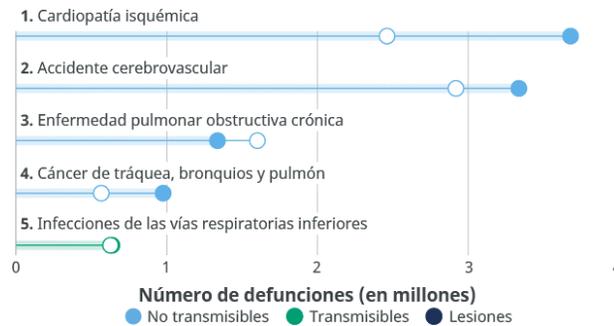


Figura 1 Tipos de accidentes cerebrovasculares.

De acuerdo con la Organización Mundial de la Salud (OMS) los accidentes cerebrovasculares (ACV) son considerados la segunda causa de mortalidad en el mundo (Figura 2). Por lo que representan aproximadamente el 11% del total de muertes [OMS, 2020].



Fuente: elaboración propia

Figura 2 Principales causas de defunción en el mundo.

En México se registran cerca de 170 mil infartos cerebrales al año. Por lo que es considerada la séptima causa de muerte en población en general al ocasionar 37 mil 453 decesos, la mayoría en hombres mayores de 65 años. Es importante mencionar que entre más rápido reciba atención médica hay menos daño, cada 15 minutos ahorrados representan el cuatro por ciento de aumento en las probabilidades de sobrevivir a la afección; es decir, las primeras horas son determinantes para una mejor recuperación o incluso salvar la vida [Gobierno de México, 2022].

El diagnóstico de un accidente cerebrovascular se basa principalmente en los síntomas (dolor de cabeza repentino que aumenta al acostarse, cambios en lucidez mental, audición, torpeza, entre otros) [Gobierno de México, 2015], pero también se realizan pruebas de diagnóstico por imagen como la resonancia magnética (RM) y la tomografía computarizada (TC). La RM es una técnica médica de obtención de imágenes que consiste en magnetizar los núcleos de los átomos del cuerpo, mediante un campo magnético muy intenso [Le Bihan, 2015].

Por un lado, las imágenes de RM por recuperación de inversión atenuada por fluido (FLAIR) son las más eficientes en el diagnóstico de ictus hiperagudo. Por otra parte, las imágenes de RM por Imagen ponderada en difusión (DWI) son altamente sensibles a la isquemia cerebral temprana, siendo su mayor utilidad dentro de las primeras seis horas posteriores al inicio del accidente cerebrovascular [Perkins, 2001].

Adicionalmente, las imágenes de RM se dividen en tres planos, el plano axial es el plano X-Y; el plano coronal es el plano X-Z y el plano sagital es el plano Y-Z. [Sriramakrishnan, 2019]. La detección asistida por computadora (CAD) es una tecnología diseñada para reducir los errores de observación, disminuir los tiempos de detección de enfermedades y, por tanto, mejorar la calidad general del diagnóstico y pronóstico médico [Castellino, 2005].

Tradicionalmente, las técnicas propuestas por Haralick eran el estándar para identificar las regiones infartadas, donde características como homogeneidad, energía, entropía y correlación se extraían de la Región de Interés (ROI) usando la Matriz de Co-ocurrencia de Niveles de Gris (GLCM) para la extracción de características de textura y para segmentación. Por una parte, [Hema, 2013] aplicó con éxito estas técnicas en cortes axiales de TC, optimizando la detección de accidentes cerebrovasculares isquémicos en un conjunto de datos de 15 casos. Posteriormente, [Revanth, 2018] fusionó estas características texturales con técnicas de clasificación contemporáneas para evaluar regiones anormales (como derrames cerebrales) en imágenes de RM. Específicamente, la investigación se benefició de la Optimización de Forrajeo Bacteriano (BFO) y la Segmentación Basada en Contorno Activo (LAC) para refinar sus resultados, basándose en un

conjunto de datos de Radiopedia con 40 imágenes en total. En un enfoque parecido, [Subudhi, 2018] empleó la GLCM y estrategias de segmentación como el filtro guiado y la transformación de watershed en imágenes de RM, para detectar y delinear con precisión los límites de las lesiones causadas por un accidente cerebrovascular utilizando 142 imágenes.

Los métodos handcraft, a pesar de su eficacia, enfrentan limitaciones como la necesidad de parametrización manual y una adaptabilidad restringida a nuevas variaciones de datos. En contraste, el deep learning ha revolucionado el procesamiento de imágenes. Estos modelos aprenden características automáticamente, se adaptan a grandes conjuntos de datos y ofrecen soluciones más potentes para desafíos actuales.

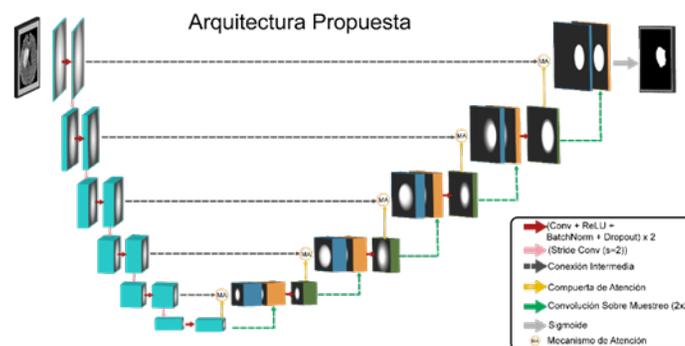
Actualmente existen diversos estudios que emplean métodos basados en aprendizaje profundo para segmentar lesiones cerebrovasculares que mejoran los resultados obtenidos por métodos tradicionales. Por ejemplo, en DRANet [Liu, 2020] se modifica la arquitectura original de U-Net, al agregar bloques residuales y un módulo de atención, por lo que se mejora el problema del desvanecimiento del gradiente y se obtienen mejores características. Se utiliza la base de datos ISLES 2015, distribuyéndola en 70% para entrenamiento y 30% para pruebas, en cuanto a la función de pérdida se utiliza Dice, Focal y Cross entropy obteniendo 76.39%, 70.92% y 62.04% respectivamente. Por otra parte, [Karthik, 2019] propone una Red Convolutiva Completa (FCN) profunda supervisada para segmentar la lesión isquémica utilizando Leaky Rectified Linear Unit (Leaky ReLU). Para evaluar el rendimiento del sistema se utilizó la base de datos ISLES 2015 destinando el 80% para entrenamiento y 20% para pruebas. Al utilizar Leaky ReLU se mejoran las características de la U-Net original obteniendo un 70% de F1 Score.

En el presente trabajo se propone un método de segmentación automática de lesiones cerebrovasculares isquémicas robusto y eficiente. Utilizando las imágenes FLAIR de la base de datos ISLES 2015 para la obtención de características que permitan la detección y localización de la zona en donde se encuentra la lesión del paciente. El algoritmo se evalúa con siete métricas, diferentes funciones de pérdida (Dice, Focal, GDFL) y distintos optimizadores (SGD y AdamW) demostrando un

buen desempeño general. Las mejoras del método propuesto a uno convencional es que se utilizan técnicas de aprendizaje profundo correctamente, así como, funciones de pérdida novedosas.

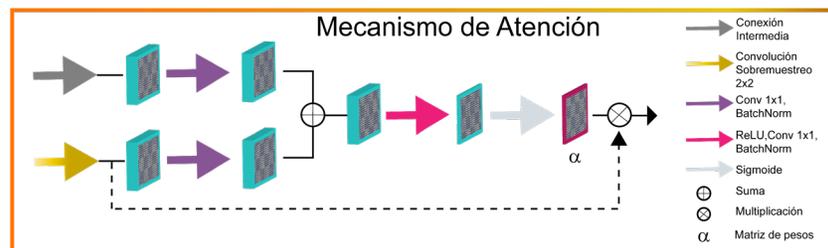
2. Métodos

En este trabajo se ha desarrollado un sistema automático de segmentación de accidentes cerebrovasculares en imágenes de RM. La arquitectura propuesta se muestra en la figura 3, dentro de la arquitectura se utilizó Attention U-Net con cinco capas convolucionales, comenzado con 32 filtros hasta 512, un *stride convolution* con valor de dos para la etapa de *downsampling*, una conexión intermedia, un mecanismo de atención (Figura 4).



Fuente: elaboración propia

Figura 3 Arquitectura basada en Attention U-Net.



Fuente: elaboración propia

Figura 4 Esquema de atención utilizado.

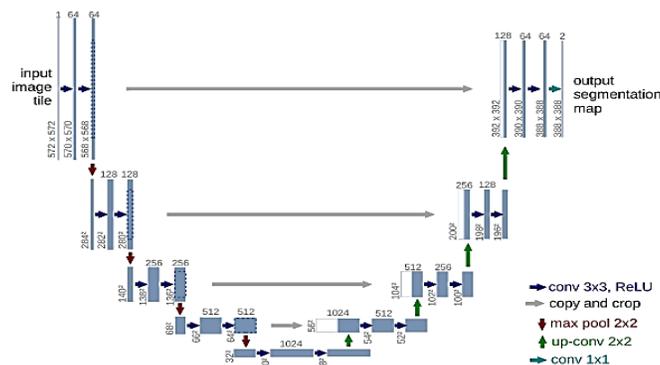
Como parte de la metodología propuesta se realizó una comparación de la arquitectura propuesta ante U-Net, además de esto se realizaron pruebas con diferentes funciones de pérdida y optimizadores para encontrar el resultado óptimo.

U-Net

U-Net es una arquitectura convolucional diseñada para la segmentación de imágenes médicas [Ronneberger, 2015]. En donde a partir de la convolución y concatenación de filtros se obtienen características relevantes de la Región de Interés (ROI). La convolución está definida por la ecuación 1, en donde I representa la imagen y K simboliza el Kernel o filtro.

$$W(i, j) = (K * I)(i, j) = \sum_m \sum_n I(i - m, j - n)K(m, n) \quad (1)$$

Esta arquitectura consta de tres partes principales: *Encoder*, *Decoder* y Conexiones Intermedias. Por una parte, el *Encoder*, extrae características semánticamente significativas de la ROI; estas características son extraídas mediante la aplicación de filtros convolucionales y posteriormente se utiliza *downsampling* a través de la operación *max pooling*, que se encarga de disminuir a la mitad el tamaño de la imagen resaltando los píxeles más relevantes en ese filtro. Por otra parte, el *Decoder*, reconstruye la imagen a su tamaño original; mientras que las conexiones intermedias ayudan a preservar la información espacial entre *Encoder-Decoder* (Figura 5).



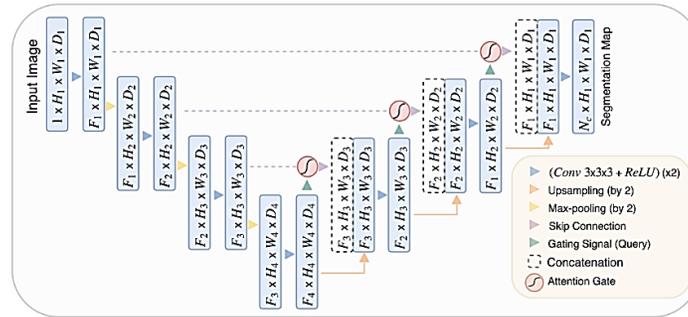
Fuente: elaboración propia

Figura 5 Arquitectura U-Net.

Attention U-Net

Attention U-Net [Oktay, 2018], es una mejora a la arquitectura base de U-Net, debido a la incorporación de un mecanismo de atención entre las conexiones intermedias llamado "Attention Gate" (AG). A diferencia de las arquitecturas U-Net

tradicionales, al contar con un AG se suprimen progresivamente las respuestas de características en regiones de fondo irrelevantes. Mejorando la segmentación general, su arquitectura se muestra en la figura 6.



Fuente: elaboración propia

Figura 6 Arquitectura Attention U-Net.

Para los modelos propuestos se utilizó la función de activación *ReLU*, la cual se define mediante la ecuación 2.

$$ReLU(x) = \begin{cases} 0, & \text{para } x < 0 \\ x, & \text{para } x \geq 0 \end{cases} \quad (2)$$

Función de pérdida

La función de pérdida juega un papel fundamental en la optimización y convergencia de los modelos. Para esta investigación se realizaron pruebas con tres diferentes tipos de funciones de pérdidas (Distribución, Región y Compuesta). Dice Loss [Yeung, 2022], es un tipo de función de pérdida basada en la región, en la cual se busca minimizar el desajuste o maximizar el traslape morfológico entre la ROI y la predicción (Ecuación 3), en donde *DSC* equivale al Coeficiente Dice-Sorensen.

$$\mathcal{L}_{DSC} = 1 - DSC \quad (3)$$

Por su parte, Focal Loss [Lin, 2008], es una generalización de la pérdida de distribución de entropía cruzada, en la cual gracias a su factor “ γ ” se pondera la aportación de los pronósticos sencillos, maximizando el enfoque en las imágenes difíciles de clasificar como se puede ver en la ecuación 4, en donde p_t significa la probabilidad de clase verdadera predicha por el modelo.

$$FL(p_t) = -(1 - p_t)^y \log(p_t) \quad (4)$$

Adicionalmente, se utilizó la función de pérdida compuesta Generalized Dice Focal Loss (GDFL) [Yeung, 2022]. La cual combina las ventajas de las funciones de pérdida basadas en la región como las basadas en la distribución, mejorando la segmentación semántica. *GDFL* se describe en la ecuación 5, donde α y β , son los coeficientes de importancia entre las *Focal Loss* y *Dice Loss*.

$$GDFL = \alpha FL + \beta \mathcal{L}_{DSC} \quad (5)$$

Optimizadores

Los optimizadores pueden explicarse como una función matemática para modificar los pesos de la red dados los gradientes, donde la elección del optimizador puede influir drásticamente en el rendimiento del modelo [Choi, 2019]. Como parte importante de la investigación se comparó el desempeño de dos optimizadores.

Por una parte, *Stochastic Gradient Descent* (SGD) el cual es un método de optimización de la tasa de aprendizaje sencillo, ecuación 6; donde θ_n son los parámetros, η es el *learning rate* y $\nabla J(\theta)$ es el gradiente. Este método suele converger a soluciones óptimas, pero con un tiempo de entrenamiento mayor.

$$\theta_{n+1} := \theta_n - \eta \nabla J(\theta) \quad (6)$$

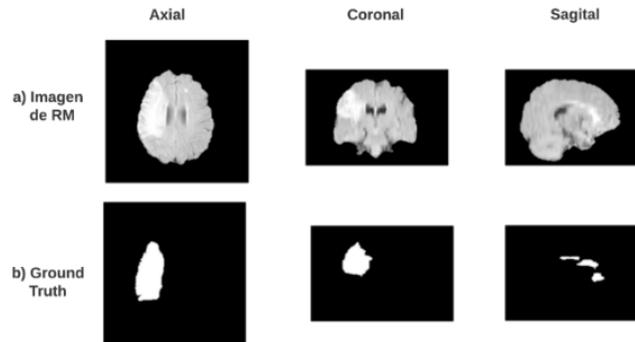
Análogamente, se utilizó AdamW (ver ecuación 7) el cual es un método más complejo que suele converger más rápido que SGD, pero normalmente a costa de una menor capacidad de generalización. Donde; $\alpha \widehat{m}_n$ y \widehat{v}_n representan el primer momento y segundo momento de los gradientes, respectivamente y λ es el *weight decay*.

$$\theta_{n+1} = \theta_n - \eta \left(\frac{\alpha \widehat{m}_n}{\sqrt{\widehat{v}_n + \epsilon}} + \lambda \theta_n \right) \quad (7)$$

Base de datos

ISLES 2015 [Maier, 2018]: Se compone de 28 pacientes y diferentes formatos de imágenes (FLAIR, DWI, T1 y T2). Con un total de imágenes con *Ground Truth* (GT) de: 1,356 para el corte axial; 1,701 para el corte coronal; 1,018 para el corte sagital.

El total de imágenes de cada corte fue distribuido en 70% para entrenamiento (80% entrenamiento y 20% validación) y 30% para pruebas. Las imágenes de los diferentes cortes que contiene la base de datos se pueden visualizar en la figura 7.



Fuente: elaboración propia

Figura 7 Imágenes de los cortes axial, coronal y sagital vs GT.

Métricas: Para la validación del rendimiento del sistema propuesto, se utilizaron las métricas de evaluación: Exactitud (ACC), Precisión (PRE), Sensibilidad (REC), Especificidad ($SPEC$), Coeficiente Dice ($F1 - Score$), Intersección sobre Unión (IoU) y Distancia Hausdorff ($h\{A,B\}$) donde: Verdaderos Positivos (TP), Verdaderos Negativos (TN), Falsos Positivos (FP) y Falsos Negativos (FN) [Müller, 2022].

Exactitud: También conocido como *Accuracy*, mide con qué frecuencia el clasificador predice correctamente. Podemos definir la precisión como la relación entre el número de predicciones correctas y el número total de predicciones representado en la ecuación 8.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

Precisión: Explica cuántos de los casos pronosticados correctamente resultaron ser positivos representada en la ecuación 9.

$$PRE = \frac{TP}{TP + FP} \quad (9)$$

Sensibilidad: También conocida como *Recall*, se utiliza para indicar el número de elementos positivos que se clasifican correctamente definida por la ecuación 10.

$$REC = \frac{TP}{TP + FN} \quad (10)$$

Especificidad: Mide la proporción de verdaderos negativos identificados con éxito por el modelo la cual está representada por la ecuación 11.

$$SPEC = \frac{TN}{TN + FP} \quad (11)$$

F1-Score: Se define como la media armónica entre la precisión y la sensibilidad y se representa en la ecuación 12.

$$F1 - Score = \frac{2TP}{2TP + FP + FN} \quad (12)$$

Intersección sobre Unión (IoU): Evalúa el solapamiento entre el GT y la región de predicción, se define por la ecuación 13.

$$IoU = \frac{TP}{TP + FP + FN} \quad (13)$$

Distancia de Hausdorff: Representa una medida de la distancia espacial entre dos conjuntos de puntos. La función $h\{A, B\}$ se denomina distancia de Hausdorff dirigida de A a B . Clasifica cada punto de A en función de su distancia al punto más cercano de B . La mayor de estas distancias determina el valor de $h\{A, B\}$ [Beauchemin, 1998], ecuación 14.

$$h\{A, B\} = \max\{\min\{d(a, b)\}\} \quad (14)$$

Ambiente de desarrollo

El modelo descrito fue implementado en una PC con las siguientes características: procesador Intel Core i7-12700K, con 32GB de RAM, GPU Nvidia 3070, un sistema operativo Linux 64-bit, Python 3.10, NumPy 1.22.4, Pandas 1.5.3 y Pytorch 2.0.1. También se realizaron pruebas con el ambiente de desarrollo de Google Colab, con una GPU Tesla T4.

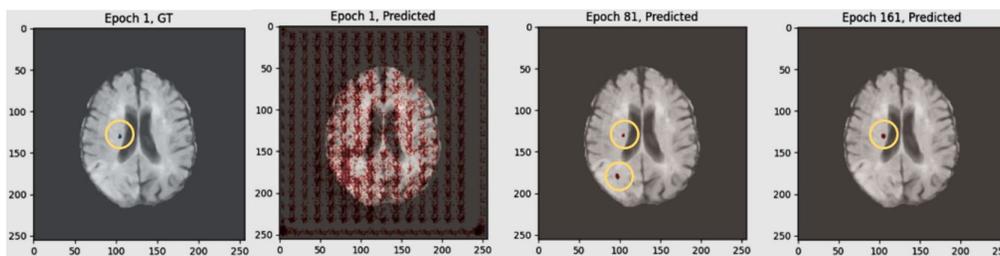
3. Resultados

En esta sección se muestran los resultados experimentales de segmentación, tanto cualitativos o perceptuales como los cuantitativos basados en las siete

métricas descritas anteriormente. Las arquitecturas propuestas se implementaron con los siguientes hiperparámetros: La función de pérdida GDFL con $\gamma = 1$, $\alpha = 0.4$ y $\beta = 0.6$, un *learning rate* de $1e - 3$. Finalmente, un *weight decay* de $1e - 2$. En el caso de *Focal Loss* se utiliza $\gamma = 1$. Por su parte las características de aumentos de datos *on the fly* fueron las siguientes: *Vertical Flip*, *Horizontal Flip*, *Random Rotate*, *Random Gamma*, *CLAHE*, *Transpose* y *Gaussian Noise* con $\text{media} = 0$, $\text{var} = 10$, todas ellas con probabilidad $p = 0.5$.

Cualitativos

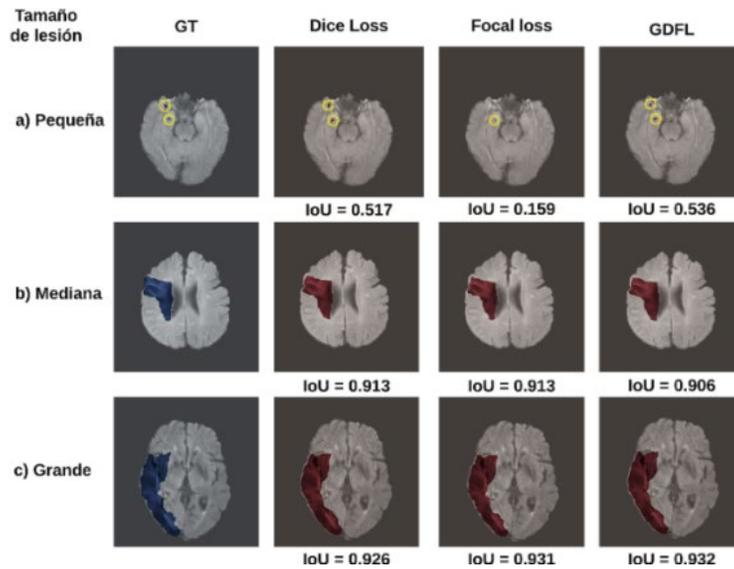
En la figura 8 se observa el aprendizaje de la red a través de las épocas donde el GT se muestra en azul y la segmentación predicha se muestra en rojo. Se puede visualizar la mejora de la ROI predicha conforme avanzan las épocas, donde en la época 1 no se han aprendido los patrones característicos de la lesión, mientras que la época 81 muestra una segmentación mucho más detallada, pero con un falso positivo, y finalmente en la época 161 se percibe un refinamiento en la predicción.



Fuente: elaboración propia

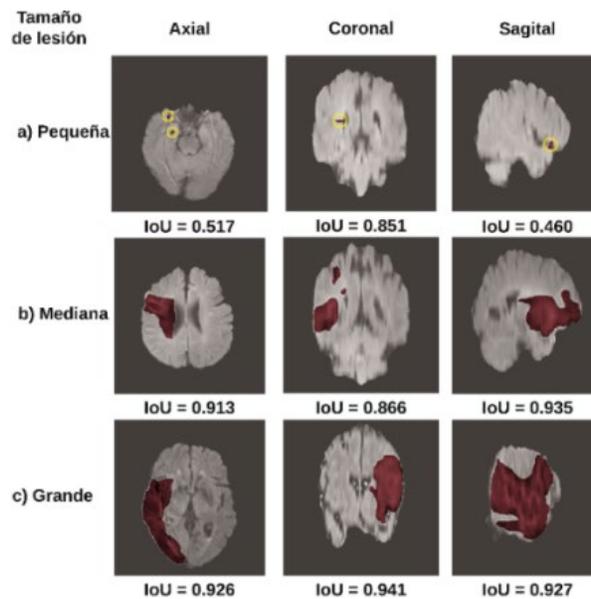
Figura 8 Aprendizaje de la red por época.

En cambio, la figura 9 muestra la comparación de las funciones de pérdida con diferentes tamaños de lesión utilizando el corte axial. Donde el mejor resultado de *IoU* lo proporciona la función de pérdida GDFL para cortes grandes y pequeños. Del mismo modo, la figura 10 muestra ejemplos de diferentes tamaños de lesiones para el corte axial, coronal y sagital con su valor de *IoU*, en el caso de lesiones de tamaño medianas y grandes se tiene un valor de *IoU* arriba de 0.90. Mientras que para las lesiones pequeñas se obtiene un *IoU* mucho menor debido a la poca cantidad de píxeles y dificultad de localización.



Fuente: elaboración propia

Figura 9 Comparación de funciones de pérdida con diferentes tamaños de lesión.



Fuente: elaboración propia

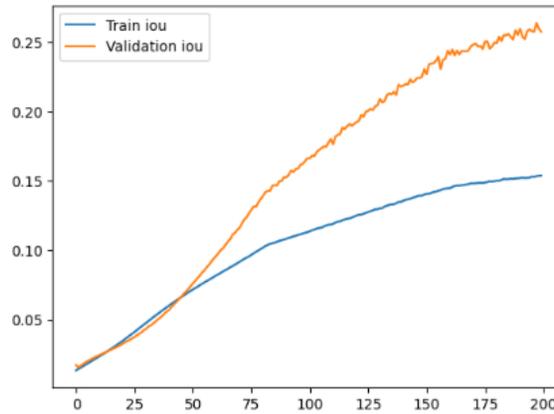
Figura 10 Diferentes tamaños de lesiones para cada corte.

Cuantitativos

En esta sección se muestran los resultados obtenidos al evaluar el método con las métricas propuestas, resaltando en negritas los mejores resultados.

En la tabla 1 se muestra una comparación entre los dos optimizadores más usados en la actualidad. Donde se puede observar que el mejor optimizador para esta tarea

en específico es AdamW puesto que para las mismas épocas de entrenamiento SGD tarda en converger, como se puede visualizar en la figura 11.



Fuente: elaboración propia

Figura 11 Curva de IoU de entrenamiento y validación para SGD.

Tabla 1 Comparación del método propuesto con GDFL para el corte axial por 200 épocas.

Optimizador	IoU	F1-Score	ACC	HD	PRE	REC	SPEC
AdamW	0.796	0.868	0.998	2.141	0.882	0.878	0.999
SGD	0.115	0.186	0.985	4.041	0.535	0.132	0.992

Fuente: elaboración propia

De forma similar la tabla 2 muestra la comparación de las arquitecturas (U-Net y el método propuesto) utilizando el mismo optimizador y la misma función de pérdida, AdamW y GDFL, respectivamente. De esta tabla se observan las ventajas de implementar mecanismos de atención a la arquitectura base, donde el método propuesto mejora en todas las métricas.

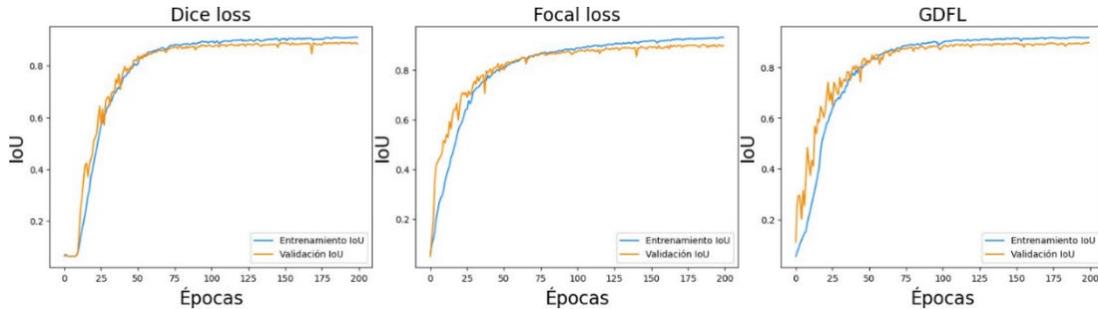
Tabla 2 Comparación de arquitecturas con GDFL, AdamW para el corte axial.

Arquitectura	IoU	F1-Score	ACC	HD	PRE	REC	SPEC
U-Net original	0.764	0.836	0.998	2.234	0.865	0.829	0.999
Método propuesto	0.796	0.868	0.998	2.141	0.882	0.878	0.999

Fuente: elaboración propia

Por su parte, la tabla 3 proporciona una comparación con diferentes funciones de pérdida, utilizando el optimizador AdamW y la arquitectura propuesta, donde los mejores resultados se obtienen con GDFL y Focal. Además de esto, se comparten

las curvas de *IoU* de entrenamiento y validación, de estas se puede observar que en ninguno de los modelos se cae en *overfitting*, ni en *underfitting* (Figura 12).



Fuente: elaboración propia

Figura 12 Curvas de *IoU* de entrenamiento y validación para cada pérdida.

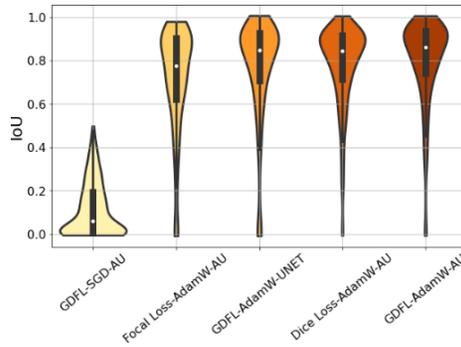
Tabla 3 Comparación del método propuesto con diferentes funciones de pérdida.

Función de pérdida	<i>IoU</i>	F1-Score	ACC	HD	PRE	REC	SPEC
Dice	0.781	0.861	0.998	2.260	0.855	0.895	0.999
Focal	0.712	0.800	0.998	2.290	0.916	0.739	0.999
GDFL	0.796	0.868	0.998	2.141	0.882	0.878	0.999

Fuente: elaboración propia

Por otra parte, las gráficas de violín muestran la función de densidad de probabilidad (PDF) de los experimentos, este tipo de gráficas proporcionan tanto una evaluación de la distribución de los datos como una inferencia estadística visualmente a simple vista [Allen, 2019]. Dicho lo anterior, la figura 13 se interpreta como la distribución de probabilidad de *IoU* para el *set* de pruebas, con la mediana señalada con un punto blanco. De acuerdo con lo que se observa de dicha figura, queda en evidencia que el modelo que máxima la probabilidad de obtener un mayor valor de *IoU* es el modelo propuesto. Así mismo, la tabla 4 muestra los resultados del modelo propuesto para cada corte (Axial, Coronal y Sagital). De donde se puede observar un rendimiento estable para los diferentes cortes, obteniendo un valor de *IoU* promedio de 0.771.

Por otra parte, en la figura 14 se visualizan los resultados del método propuesto para cada corte en una gráfica de violín. De esta gráfica se percibe que la distribución de probabilidad de *IoU* para cada corte está correctamente desplazada hacia valores altos con una mediana arriba de 0.800 en cada caso.



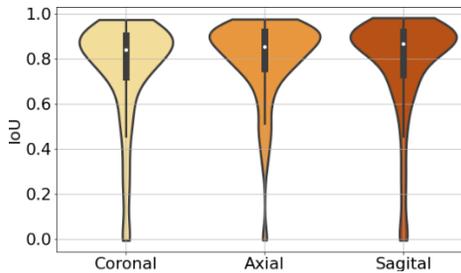
Fuente: elaboración propia

Figura 13 Comparación de funciones de pérdida y arquitecturas utilizadas.

Tabla 4 Resultados del método propuesto en diferentes cortes.

Corte	IoU	F1-Score	ACC	HD	PRE	REC	SPEC
Axial	0.791	0.869	0.998	2.196	0.871	0.89	0.999
Coronal	0.774	0.845	0.998	2.260	0.874	0.848	0.999
Sagital	0.749	0.828	0.995	2.942	0.832	0.851	0.997
Promedio	0.771	0.847	0.997	2.466	0.859	0.863	0.998

Fuente: elaboración propia



Fuente: elaboración propia

Figura 14 Resultados del método propuesto para los cortes corona, axial y sagital.

Finalmente, la tabla 5 expone una comparación del método propuesto con los del estado del arte. Se puede observar que el método propuesto obtiene resultados que mejoran en todas las métricas, en especial en F1-Score donde mejora por un 10%, ante el segundo mejor DRANet.

Tabla 5 Comparación del método propuesto con el estado del arte.

Método	IoU	F1-Score	ACC	HD	PRE	REC	SPEC
DRANet [Liu, 2020]	-	0.762	-	3.36	-	-	-
FCN [Karthik, 2019]	-	0.700	-	-	-	-	-
Método propuesto	0.796	0.868	0.998	2.141	0.882	0.878	0.999

Fuente: elaboración propia

4. Discusión

Se propuso un método con una variante modificada de Attention U-Net, donde se comprueba que los mejores resultados son con el optimizador AdamW, puesto que durante 200 épocas converge mucho más rápido que SGD. Adicionalmente, al comparar el método propuesto ante el estado del arte ([Liu, 2020] y [Karthik, 2019]) se observa que las funciones de pérdida compuestas pueden hacer un cambio significativo en la *IoU* al tomar lo mejor de las pérdidas de región y de distribución. Además, se demuestra que la arquitectura propuesta basada en Attention U-Net mejora U-Net gracias a los mecanismos de atención.

Finalmente, la investigación resalta la importancia de la elección de los hiperparámetros como funciones de pérdida y optimizadores para la segmentación semántica. A pesar de los resultados prometedores obtenidos en esta investigación todavía existe un campo de mejora, donde se podría obtener una clasificación más robusta al entrenar la red con diferentes bases de datos y con diferentes modalidades de imágenes como podría ser FLAIR, DWI, T1, T2 y ADC.

5. Conclusiones

En el presente trabajo se realizó la implementación de una arquitectura convolucional basada en Attention U-Net para la segmentación de imágenes de resonancia magnética de la base de datos ISLES 2015, para FLAIR y con los cortes Axial, Coronal y Sagital. Por una parte, se comparan los optimizadores y funciones de pérdida que maximizan la *IoU*, donde queda demostrado que los mejores resultados se obtienen con GDFL, AdamW y la arquitectura propuesta.

Los resultados se comparan con siete métricas y ante dos de los mejores métodos consultados en el estado del arte para esta base de datos, obteniendo resultados correctos para cada corte, mejorando la segmentación de la región de interés. Adicionalmente se comparten las curvas de entrenamiento y los diagramas de violín que demuestran el correcto desempeño del modelo propuesto. Por lo que este método puede ser utilizado como un método complementario de visualización y de segmentación del área de la lesión que permita mejorar el diagnóstico médico.

Como trabajo futuro se considera extender la investigación para analizar diferentes tipos de imágenes como FLAIR, DWI, T1, T2 y ADC. Además de analizar cual combinación de tipos de imágenes facilita la segmentación.

6. Bibliografía y Referencias

- [1] Allen, M., Poggiali, D., Whitaker, K., Marshall, TR., van Langen, J., & Kievit, RA. Raincloud plots: a multi-platform tool for robust data visualization. *Wellcome Open Research*. 2019; 4:63. DOI: 10.12688/wellcomeopenres.15191.2. PMID: 31069261; PMCID: PMC6480976.
- [2] Castellino, R. A. Computer aided detection (CAD): an overview. *Cancer Imaging*, 2005 Aug 23; 5(1):17-9. doi: 10.1102/1470-7330.2005.0018. PMID: 16154813; PMCID: PMC1665219.
- [3] Che-Wei, L., and Jeen-Shing, W. A digital circuit design of hyperbolic tangent sigmoid function for neural networks, 2008 IEEE International Symposium on Circuits and Systems (ISCAS), Seattle, WA, USA, 2008, pp. 856-859, doi: 10.1109/ISCAS.2008.4541553.
- [4] Choi, D., Shallue, C. J., Nado, Z., Lee, J., Maddison, C. J., & Dahl, G. E., (2019). On Empirical Comparisons of Optimizers for Deep Learning. 1. <http://arxiv.org/abs/1910.05446>.
- [5] Gobierno de México, (2015). Accidentes Cerebrovasculares. Recuperado de <https://n9.cl/ergng>.
- [6] Gobierno de México, (2022). En 2021, ictus o enfermedad vascular cerebral ocasionó más de 37 mil decesos en México. Recuperado de <https://n9.cl/r8z64>.
- [7] Hema-Rajini, N., & Bhavani, R., (2013). Computer aided detection of ischemic stroke using segmentation and texture features. *Measurement: Journal of the International Measurement Confederation*, 46(6), 1865–1874. <https://doi.org/10.1016/j.measurement.2013.01.010>.
- [8] IMSS, (2015). Enfermedad Vascular Cerebral. IMSS. Recuperado de <https://n9.cl/mlwn3>.

- [9] K. Revanth, T. D. Varsha shree, N. S. Madhava Raja and V. Rajinikanth, Computational Investigation of Stroke Lesion Segmentation from Flair/DW Modality MRI, 2018 Fourth International Conference on Biosignals, Images and Instrumentation (ICBSII), Chennai, India, 2018, pp. 206-212, doi: 10.1109/ICBSII.2018.8524617.
- [10] Karthik, R., Gupta, U., Jha, A., Rajalakshmi, R., & Menaka, R., (2019). A deep supervised approach for ischemic lesion segmentation from multimodal MRI using Fully Convolutional Network. *Applied Soft Computing Journal*, 84(105685). <https://doi.org/10.1016/j.asoc.2019.105685>.
- [11] Le Bihan, D., & Lima, M., (2015). Diffusion magnetic resonance imaging: What water tells us about biological tissues. *PLoS Biology*, 13(7), e1002203. <https://doi.org/10.1371/journal.pbio.1002203>.
- [12] Liu, L., Kurgan, L., Wu, F. X., & Wang, J., (2020). Attention convolutional neural network for accurate segmentation and quantification of lesions in ischemic stroke disease. *Medical Image Analysis*, 65(101791), 101791. <https://doi.org/10.1016/j.media.2020.101791>.
- [13] M. Beauchemin, K.P.B. Thomson & G. Edwards, (1998) On the Hausdorff Distance Used for the Evaluation of Segmentation Results, *Canadian Journal of Remote Sensing*, 24:1, 3-8, DOI: 10.1080/07038992.1998.10874685.
- [14] Maier, O., Menze, B. H., von der Gablentz, J., Häni, L., Heinrich, M. P., Liebrand, M., Winzeck, S., Basit, A., Bentley, P., Chen, L., Christiaens, D., Dutil, F., Egger, K., Feng, C., Glocker, B., Götz, M., Haeck, T., Halme, H.-L., Havaei, M., Reyes, M., (2017). ISLES 2015 - A public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. *Medical Image Analysis*, 35, 250–269. <https://doi.org/10.1016/j.media.2016.07.009>.
- [15] MemorialCare, (2023). Stroke. MemorialCare. Recuperado de <https://www.memorialcare.org/services/stroke-care/stroke>.
- [16] Müller, D., Soto-Rey, I. & Kramer, F. Towards a guideline for evaluation metrics in medical image segmentation. *BMC Res Notes* 15, 210, 2022. <https://doi.org/10.1186/s13104-022-06096-y>.

- [17] Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N. Y., Kainz, B., Glocker, B., & Rueckert, D., (2018). Attention U-Net: Learning where to look for the pancreas. In arXiv [cs.CV]. <http://arxiv.org/abs/1804.03999>.
- [18] OMS, (2020). Las 10 principales causas de defunción. OMS. recuperado de <https://n9.cl/07m39>.
- [19] Perkins, C. J., Kahya, E., Roque, C. T., Roche, P. E., & Newman, G. C., (2001). Fluid-attenuated inversion recovery and Diffusion- and perfusion-weighted MRI abnormalities in 117 consecutive patients with stroke symptoms. *Stroke; a Journal of Cerebral Circulation*, 32(12), 2774–2781. <https://doi.org/10.1161/hs1201.099634>.
- [20] Ronneberger, O., Fischer, P. & Brox, T., (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In N. Navab, J. Hornegger, W. M. Wells & A. F. Frangi (eds.), *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III* (pp. 234--241). Springer International Publishing. ISBN: 978-3-319-24574-4.
- [21] Sriramakrishnan, P., Kalaiselvi, T., Thirumalaiselvi, M., Padmapriya, S. T., & Ramkumar, S., (2019). A Role of Medical Imaging Techniques in Human Brain Tumor Treatment. *International Journal of Recent Technology and Engineering*, 8(4S2), 565–568. <https://n9.cl/glu2f>.
- [22] Subudhi, A., Jena, S. & Sabut, S. Delineation of the ischemic stroke lesion based on watershed and relative fuzzy connectedness in brain MRI. *Med Biol Eng Comput* 56, 795–807, 2018. <https://n9.cl/0zgdg>.
- [23] Yeung, M., Sala, E., Schönlieb, C. B., & Rundo, L., (2022). Unified Focal loss: Generalising Dice and cross entropy-based losses to handle class imbalanced medical image segmentation. *Computerized Medical Imaging and Graphics: The Official Journal of the Computerized Medical Imaging Society*, 95(102026), 102026. <https://doi.org/10.1016/j.compmedimag.2021.102026>.