



**MODELO MULTIMODAL PARA LA SEGMENTACIÓN DE LESIONES DE  
ACV ISQUÉMICO EN ESTUDIOS DE RESONANCIA MAGNÉTICA**

**JULIAN ESTEBAN GARCIA DUARTE**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FISICOMECAICAS  
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA  
PROGRAMA DE PREGRADO EN INGENIERÍA DE SISTEMAS E  
INFORMÁTICA  
BUCARAMANGA**

**2024**

**MODELO MULTIMODAL PARA LA SEGMENTACIÓN DE LESIONES DE  
ACV ISQUÉMICO EN ESTUDIOS DE RESONANCIA MAGNÉTICA**

**JULIÁN ESTEBAN GARCÍA DUARTE**

**Trabajo de Grado para optar por el título de:  
Ingeniero de Sistemas**

**Director:**

**Fabio Martínez Carrillo**

**Ph.D in Systems and Computer Engineering**

**Codirector:**

**Santiago Gómez Hernández**

**M.Sc in Systems and Computer Engineering**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FISICOMECHANICAS  
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA  
PROGRAMA DE PREGRADO EN INGENIERÍA DE SISTEMAS E  
INFORMÁTICA  
BUCARAMANGA**

**2024**

## AGRADECIMIENTOS

El autor expresa sus agradecimientos:

Primero que todo, al Profesor Fabio Martínez, director del grupo de investigación BIVL2ab por su guía durante el desarrollo de este proyecto. Su experiencia y conocimientos en el campo de la inteligencia artificial fueron fundamentales para el éxito de mi proyecto. También quiero agradecer a Santiago Gómez, que brindó un apoyo constante y valioso en el desarrollo del proyecto. Su experiencia en el área de trabajo y trayectoria académica fueron clave para llevar a cabo la implementación de mi trabajo.

Quiero agradecer a mi familia por el apoyo incondicional que me han brindado durante toda mi carrera y en especial durante la elaboración de este proyecto. Su confianza en mis capacidades ha sido fundamental para alcanzar los objetivos que me he propuesto. Agradezco a mis amigos por haberme acompañado y alentado durante todo este tiempo. Su ánimo y disposición a escuchar mis ideas y reflexiones han sido de gran ayuda para mi crecimiento personal y profesional. También, quiero expresar mi agradecimiento a la Escuela de Ingeniería de Sistemas por brindarme una formación de calidad y completa durante estos años de estudio. Los conocimientos adquiridos en esta institución me han permitido enfrentar con éxito los desafíos que se me han presentado a lo largo de mi proyecto.

## CONTENIDO

	page
<b>INTRODUCCIÓN</b> . . . . .	<b>10</b>
<b>1. FUNDAMENTOS Y TRABAJO PREVIO</b> . . . . .	<b>13</b>
1.1. Accidente Cerebrovascular Isquémico . . . . .	13
1.1.1 Imágenes de Resonancia Magnética . . . . .	14
1.2. U-Net: una arquitectura para la segmentación de imágenes médicas . . . . .	17
1.3. Estrategias para segmentar lesiones de ACV isquémico . . . . .	19
<b>2. PROBLEMA DE INVESTIGACIÓN</b> . . . . .	<b>23</b>
<b>3. OBJETIVOS</b> . . . . .	<b>24</b>
<b>4. MÉTODO PROPUESTO</b> . . . . .	<b>25</b>
4.1. Modelo multimodal y multisegmentacion . . . . .	25
4.2. Fusión temprana del FLAIR . . . . .	29
4.3. Fusión tardía del FLAIR . . . . .	30
<b>5. CONFIGURACIÓN EXPERIMENTAL</b> . . . . .	<b>32</b>
5.1. Datos . . . . .	32
5.2. Configuración de la arquitectura . . . . .	33
5.3. Métricas de evaluación . . . . .	33
5.4. Evaluación del desempeño en la segmentación de lesiones de ACV isquémico . . . . .	36
<b>6. EVALUACIÓN Y RESULTADOS</b> . . . . .	<b>38</b>
6.1. Evaluación de arquitecturas unimodales . . . . .	38

6.2. Evaluación multimodal con U-Net clásica . . . . .	40
6.3. Evaluación multimodal con base en la Dual U-Net y la Fusión temprana . . . . .	42
6.4. Evaluación multimodal con base en la Dual U-Net y la Fusión tardía . . . . .	44
<b>7. CONCLUSIONES Y TRABAJO FUTURO . . . . .</b>	<b>46</b>
<b>BIBLIOGRAFÍA . . . . .</b>	<b>48</b>

## LISTA DE FIGURAS

	<b>page</b>
Figura 1. Lesión isquémica en estudios de resonancia magnética. . . . .	15
Figura 2. Mapa de difusión isotrópico para paciente de ACV. . . . .	16
Figura 3. Coeficiente Aparente de Difusión para paciente de ACV. . . . .	16
Figura 4. Arquitectura U-Net . . . . .	18
Figura 5. Representación gráfica de los métodos de integración multimodal para la segmentación de la lesión isquémica. . . . .	26
Figura 6. Representación gráfica de la fusión temprana del FLAIR. . . . .	30
Figura 7. Representación gráfica de la fusión tardía del FLAIR. . . . .	31

## LISTA DE TABLAS

	<b>page</b>
Tabla 1. Métricas Computacionales Unimodal . . . . .	39
Tabla 2. Métricas Médicas Unimodal . . . . .	40
Tabla 3. Métricas Computacionales Multimodal U-Net Clásica . . . . .	41
Tabla 4. Métricas Médicas Multimodal U-Net Clásica . . . . .	42
Tabla 5. Métricas Computacionales Fusión Temprana Dual U-Net . . . . .	43
Tabla 6. Métricas Médicas Fusión Temprana Dual U-Net . . . . .	44
Tabla 7. Métricas Computacionales Fusión Tardía Dual U-Net . . . . .	44
Tabla 8. Métricas Médicas Fusión Tardía Dual U-Net . . . . .	45

## RESUMEN

**TÍTULO:** MODELO MULTIMODAL PARA LA SEGMENTACIÓN DE LESIONES DE ACV ISQUÉMICO EN ESTUDIOS DE RESONANCIA MAGNÉTICA \*

**AUTOR:** JULIAN ESTEBAN GARCIA DUARTE \*\*

**PALABRAS CLAVE:** REPRESENTACIÓN MULTIMODAL, ACCIDENTE CEREBROVASCULAR, COEFICIENTE APARENTE DE DIFUSIÓN, IMAGEN DE DIFUSIÓN PONDERADA.

**DESCRIPCIÓN:** Los accidentes cerebrovasculares (ACV) son la segunda causa de muerte a nivel mundial, reportando más de 100 millones de personas afectadas anualmente. Accidentes cerebrovasculares existen de dos tipos, hemorrágico e isquémico. Particularmente, el ACV isquémico es el tipo de lesión con mayor prevalencia y con un alto riesgo de morbilidad asociado. La obtención de hallazgos radiológicos para la cuantificación de la lesión en estudios de resonancia magnética es clave para el diagnóstico, establecer un procedimiento clínico y estimar un pronóstico del paciente. Esta tarea se realiza mediante estudios imagenológicos especiales (estudios paramétricos) que permiten cuantificar la hemodinámica cerebral y brindan información sobre el tejido intracraneal, como por ejemplo el Coeficiente Aparente de Difusión (ADC, por sus siglas en inglés), Imagen de Difusión Ponderada (DWI, por sus siglas en inglés), Recuperación de la Inversión Atenuada por Fluidos (FLAIR, por sus siglas en inglés), entre otras. De forma observacional, un experto analiza múltiples de estas imágenes al mismo tiempo, usando la complementariedad entre estas para realizar la segmentación de la lesión isquémica. A pesar de los avances reportados en la literatura, las estrategias computacionales siguen presentando limitaciones para capturar la complejidad y alta variabilidad de la lesión. Este trabajo busca implementar un esquema multimodal que integre al menos dos estudios paramétricos de resonancia magnética, aprovechando hallazgos radiológicos complementarios que permitan la cuantificación de las lesiones relacionadas con ACV. Para el entrenamiento y validación se contó con un conjunto de datos propio del grupo de investigación *BIVL<sup>2</sup>ab* (Biomedical Imaging, Vision and Learning Laboratory) y también se exploró la posibilidad de validaciones en conjuntos públicos disponibles. El modelo implementado fue validado según su capacidad de segmentar lesiones isquémicas para un conjunto de datos no usado en el entrenamiento.

---

\* Trabajo de investigación

\*\* Facultad de Ingenierías Físicomecánicas. Escuela de Ingeniería de Sistemas e Informática. Director: Fabio Martínez, PhD.

## ABSTRACT

**TITLE:** MULTIMODAL MODEL FOR THE SEGMENTATION OF ISCHEMIC STROKE LESIONS IN MAGNETIC RESONANCE IMAGING STUDIES. \*

**AUTHOR:** JULIAN ESTEBAN GARCIA DUARTE \*\*

**KEYWORDS:** MULTIMODAL REPRESENTATION, ISCHEMIC STROKE, APPARENT DIFFUSION COEFFICIENT, DIFFUSION-WEIGHTED IMAGE.

**DESCRIPTION:** Cerebrovascular accidents are the second leading cause of death worldwide, reporting more than 100 million people affected annually. Strokes exist in two types, hemorrhagic and ischemic. Particularly, ischemic stroke is the most prevalent type of injury with an associated high risk of morbidity. Obtaining radiological findings for the quantification of the lesion in MRI studies is key for diagnosis, establishing a clinical procedure and estimating the patient's prognosis. This task is performed by means of special imaging studies (parametric studies) that allow the quantification of cerebral hemodynamics and provide information on intracranial tissue, such as Apparent Diffusion Coefficient (ADC), Diffusion Weighted Imaging (DWI), Fluid Attenuated Inversion Recovery (FLAIR), among others. Observationally, an expert analyzes multiple of these images at the same time, using the complementarity between them to perform ischemic lesion segmentation. Despite the advances reported in the literature, computational strategies still present limitations to capture the complexity and high variability of the lesion. This work seeks to implement a multimodal scheme that integrates at least two parametric MRI studies, taking advantage of complementary radiological findings that allow quantification of stroke-related lesions. For training and validation we will rely on a dataset owned by the *BIVL<sup>2</sup>ab* (Biomedical Imaging, Vision and Learning Laboratory) research group and will also explore the possibility of validations on publicly available sets. The implemented model will be validated according to its ability to segment ischemic lesions for a dataset not used in training.

---

\* Research work

\*\* Faculty of Physics-Mechanics Engineering. School of Systems Engineering and Informatics. Advisor: Fabio Martínez, PhD.

## INTRODUCCIÓN

El accidente cerebrovascular (ACV) es la segunda causa de mortalidad y la tercera causa de muerte y discapacidad a nivel mundial. Su incidencia ha incrementado un 70%, su mortalidad un 43% y la reducción en la calidad de vida a quienes lo padecen un 143% medido DALYs<sup>1</sup>. Los efectos adversos del ACV afectan principalmente a países en vía de desarrollo y subdesarrollados (86% de las muertes y 89% de DALYs)<sup>1</sup>. El ACV es una enfermedad causada por la interrupción del flujo sanguíneo en el cerebro por la oclusión de un vaso por un coágulo de sangre (ACV isquémico trombótico), por un émbolo que se origina en otro lugar del organismo (ACV isquémico embólico) o por la ruptura de un vaso sanguíneo (ACV hemorrágico). El tipo de ACV más común es el isquémico, con alrededor de un 62% de todos los casos a nivel mundial<sup>1</sup>. El protocolo de tratamiento de pacientes con sospecha de ACV incluye la captura de estudios radiológicos para caracterizar las lesiones isquémicas. Los estudios que ofrecen mayor sensibilidad para discriminar entre tejido sano y dañado son los de resonancia magnética (MRI, por sus siglas en inglés). En estos estudios se capturan imágenes ponderadas por difusión (DWI, por sus siglas en inglés) y de coeficiente aparente de difusión (ADC, por sus siglas en inglés), capaces de resaltar los patrones ocasionados por el edema cerebral citotóxico, observado con alta frecuencia en la isquemia cerebral. Sobre estas modalidades, los radiólogos llevan a cabo delineaciones manuales para localizar las zonas hipoperfundidas y cuantificar su volumen. Este proceso es parte fundamental para informar un buen diagnóstico, procedimiento clínico y estimación del pronóstico del paciente<sup>2</sup>. Sin embargo, este procedimiento es altamente subjetivo (estudios

---

<sup>1</sup> Valery L FEIGIN et al. “World Stroke Organization (WSO): global stroke fact sheet 2022”. In: *International Journal of Stroke* 17.1 (2022), pp. 18–29.

<sup>2</sup> Seung-Hoon LEE. *Stroke Revisited: Diagnosis and treatment of ischemic stroke*. Jan. 2017. DOI: 10.1007/978-981-10-1424-6.

reportan diferencias estadísticas significativas entre marcaciones de diferentes expertos)<sup>3</sup> y tedioso (cerca de 15 minutos en cada caso)<sup>4</sup>. Actualmente, en la literatura se han propuesto arquitecturas de aprendizaje profundo de tipo codificador-decodificador para llevar a cabo la segmentación de lesiones de ACV isquémico sobre estudios de MRI. Más específicamente, se han propuesto trabajos que utilizan arquitecturas U-Net que procesan una única modalidad de resonancia magnética<sup>567</sup>. Estas arquitecturas integran conexiones de salto para reintegrar características del codificador en el decodificador y así obtener mejores segmentaciones de lesiones de ACV isquémico. Sin embargo, estos trabajos son sensibles a un incremento de falsos positivos en las predicciones, debido a artefactos como el T2 *shine-through*, que hace referencia a altas intensidades en DWI que no corresponden a restricción en la difusión del tejido, sino que es producto de lecturas altas de intensidad en la imagen T2 (debido a un tiempo prolongado de decaimiento en tejido normal) que brillan a través (*shine-through*, en inglés) de esta hasta el DWI<sup>8</sup>. Además, se han propuesto trabajos que construyeron arquitecturas multimodales para

---

<sup>3</sup> Anders B NEUMANN et al. “Interrater agreement for final infarct MRI lesion delineation”. In: *Stroke* 40.12 (2009), pp. 3768–3771.

<sup>4</sup> Anne L MARTEL et al. “Measurement of infarct volume in stroke patients using adaptive segmentation of diffusion weighted MR images”. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI’99: Second International Conference, Cambridge, UK, September 19–22, 1999. Proceedings 2*. Springer. 1999, pp. 22–31.

<sup>5</sup> Manas Kumar NAG et al. “Computer-assisted delineation of cerebral infarct from diffusion-weighted MRI using Gaussian mixture model”. In: *International Journal of Computer Assisted Radiology and Surgery* 12.4 (Jan. 2017), pp. 539–552. DOI: 10.1007/s11548-017-1520-x.

<sup>6</sup> Hyunkwang SHIN et al. “Automated segmentation of chronic stroke lesion using efficient U-Net architecture”. In: *Biocybernetics and Biomedical Engineering* 42.1 (2022), pp. 285–294. DOI: <https://doi.org/10.1016/j.bbe.2022.01.002>.

<sup>7</sup> Shuqiang WANG et al. “Brain stroke lesion segmentation using consistent perception generative adversarial network”. In: *Neural Computing and Applications* 34.11 (Mar. 2022), pp. 8657–8669. DOI: 10.1007/s00521-021-06816-8.

<sup>8</sup> Stéphane SILVERA et al. “Spontaneous Intracerebral Hematoma on Diffusion-weighted Images: Influence of T2-shine-through and T2-blackout Effects”. In: *American Journal of Neuroradiology* 26.2 (2005), pp. 236–241. eprint: <https://www.ajnr.org/content/26/2/236.full.pdf>.

segmentar lesiones de ACV desde múltiples modalidades de MRI <sup>91011</sup>. Estos trabajos tienen la ventaja de que integran la información de cada modalidad en la estimación de las lesiones de ACV isquémico. Sin embargo, llevan a cabo la estimación de las lesiones a partir de una simple concatenación de modalidades MRI, limitando la captura eficiente de los patrones visuales de las lesiones de ACV. En este trabajo se implementó un esquema de aprendizaje profundo para la segmentación de lesiones de ACV isquémico, utilizando arquitecturas codificador-decodificador que aprovechan la información tisular complementaria de modalidades multiparamétricas MRI. La arquitectura codifica patrones visuales de ACV en representaciones compactas por medio de ramas de codificación diseñadas específicamente para cada modalidad, y estima los hallazgos radiológicos de ACV en cada modalidad mediante ramas decodificadoras independientes. El esquema multi-segmentación induce el aprendizaje de características altamente correlacionadas presentes en las modalidades MRI, contribuyendo a un mejoramiento en la segmentación de las lesiones. Para el entrenamiento y validación del enfoque propuesto, se utilizó un conjunto de datos privado con 186 estudios de pacientes atendidos en una unidad especializada en ACV. Cada estudio cuenta con las modalidades ADC, DWI, y en algunos casos, FLAIR, acompañadas de anotaciones realizadas por múltiples expertos.

---

<sup>9</sup> Santiago GÓMEZ et al. “Ischemic Stroke Segmentation from a Cross-Domain Representation in Multimodal Diffusion Studies”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2023, pp. 776–785.

<sup>10</sup> Abhishek BAL et al. “A robust ischemic stroke lesion segmentation technique using two-pathway 3D deep neural network in MR images”. In: *Multimedia Tools and Applications* (Oct. 2023). DOI: 10.1007/s11042-023-16689-9.

<sup>11</sup> Rongzhao ZHANG et al. “Automatic Segmentation of Acute Ischemic Stroke From DWI Using 3-D Fully Convolutional DenseNets”. In: *IEEE Transactions on Medical Imaging* 37.9 (2018), pp. 2149–2160. DOI: 10.1109/TMI.2018.2821244.

## 1. FUNDAMENTOS Y TRABAJO PREVIO

### 1.1. Accidente Cerebrovascular Isquémico

El accidente cerebrovascular (ACV) es una lesión del tejido intracraneal producida por la interrupción de flujo sanguíneo <sup>12</sup>. Dependiendo de su causa, este puede ser clasificado como hemorrágico o isquémico. El ACV isquémico (el más frecuente con un 60% de todos los casos <sup>1</sup>) ocurre por una obstrucción de un vaso sanguíneo debido a un agente obstructor (normalmente un coágulo de sangre). Esta obstrucción limita el suministro de glucosa y oxígeno al tejido circundante, iniciando una reacción en cadena a nivel celular en donde las neuronas agotan los recursos disponibles y conlleva, finalmente, a la muerte del tejido<sup>13</sup>. Si bien todo el tejido afectado sufre una caída en la cantidad de sangre perfundida, el área afectada se puede dividir en dos partes: la región inmediatamente siguiente a la obstrucción, conocida como núcleo del infarto, y el tejido que rodea a este, conocido como penumbra (ver ilustración en la **figura 1-A**). A diferencia de la penumbra, el núcleo isquémico carece de lo necesario para sostener los procesos celulares mínimos, constituyéndolo en tejido no recuperable<sup>1415</sup>. Inmediatamente luego del ACV isquémico, los vasos sanguíneos se dilatan y rutas de respaldo entran en funcionamiento irrigando el tejido afectado. Este cambio en la hemodinámica del cerebro se da como respuesta fisiológica a la isquemia y puede retrasar los síntomas del ACV y mantener la integridad del cerebro por un periodo de tiempo corto, impidiendo que la penumbra se infarte. Luego de un periodo de

---

<sup>12</sup> Caplan's Stroke. Aug. 5, 2016. DOI: 10.1017/cbo9781316095805.

<sup>13</sup> Steven K. FESKE. "Ischemic Stroke". In: *The American Journal of Medicine* 134.12 (Dec. 2021), pp. 1457–1464. DOI: 10.1016/j.amjmed.2021.07.027.

<sup>14</sup> Trent M. WOODRUFF et al. "Pathophysiology, treatment, and animal and cellular models of human ischemic stroke". In: *Molecular Neurodegeneration* 6.1 (Jan. 2011). DOI: 10.1186/1750-1326-6-11.

<sup>15</sup> S STRANDGAARD and OB PAULSON. "Pathophysiology of stroke". In: *Journal of cardiovascular pharmacology* 15 (1990), S38–S42.

tiempo (que es diferente para cada paciente), las vías de irrigación auxiliares fallan y el cerebro se ve incapaz de mantener el nivel de perfusión, aumentando el radio del núcleo de infarto. Por lo tanto, es fundamental realizar un análisis rápido del paciente mediante imágenes diagnósticas para iniciar el tratamiento adecuado. La principal herramienta diagnóstica para realizar una caracterización de la lesión consisten en la resonancia magnética (RM) cuyos parámetros e imágenes resultantes se utilizan para localizar y cuantificar la lesión. A continuación se describe brevemente estas secuencias imagenológicas.

### 1.1.1. Imágenes de Resonancia Magnética

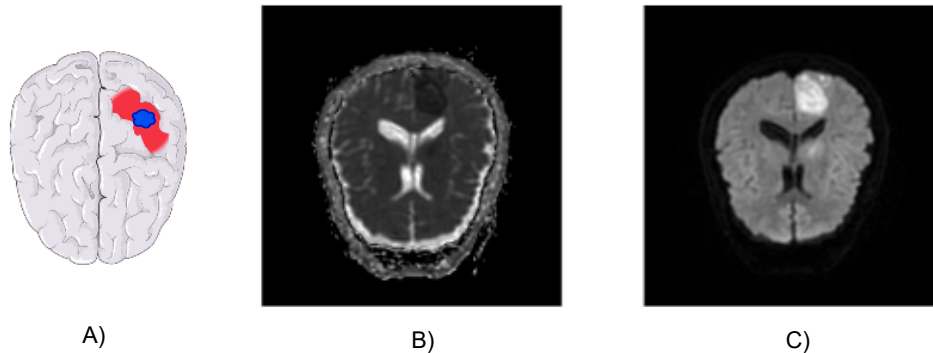
Las imágenes diagnósticas de resonancia magnética (MRI, por sus siglas en inglés) permiten describir estructural y funcionalmente el tejido intracraneal, cervical y espinal. Estas imágenes ofrecen una mayor sensibilidad para evidenciar cambios en el tejido afectado por lesiones asociadas al ACV. Además, estas técnicas de adquisición pueden estar acompañadas por agentes de contraste (imágenes de perfusión) que permiten resumir observaciones asociadas a las trayectorias descritas por un líquido sensible a la adquisición. Así, es posible lograr, por ejemplo, estimaciones del núcleo isquémico, así como también determinar los tiempos de exposición del ACV, y determinar el avance de la lesión.

También, otras modalidades multiparamétricas capturadas de un resonador magnético, pueden estimar el movimiento de las moléculas de agua, lo que tiene una alta correlación con la diferenciación entre tejido sano y tejido afectado por una lesión de ACV. Particularmente, las modalidades de ADC (coeficiente de difusión aparente) son particularmente útiles para el diagnóstico temprano de ACV isquémico, ya que pueden identificar áreas de baja perfusión en el tejido cerebral minutos después de que se produce el evento. Estas áreas indican un edema citotóxico y son una señal temprana de la lesión cerebral irreversible. A continuación se detallan las imágenes multiparamétricas de interés para el análisis de ACV.

- **Imagen de Difusión Ponderada (DWI por sus siglas en inglés):** Estas imágenes son mapas de difusión isotrópicos que reflejan el movimiento Browniano de las partículas de

## Figura 1

*Lesión isquémica en estudios de resonancia magnética*



*Nota.* A) Imagen ilustrativa del cerebro, en rojo la penumbra isquémica y en azul el núcleo isquémico. B) Lesión isquémica vista en una imagen de ADC, la lesión se puede observar como la región hipointensa (más oscura respecto al tejido del cerebro) en la parte superior derecha del cerebro. C) Lesión isquémica vista en una imagen de DWI, la lesión se puede observar como la región hiperintensa (más brillante respecto al tejido del cerebro) en la parte superior derecha del cerebro.

agua dentro del tejido. Estas modalidades se componen por la combinación de la señal T2 y los valores de difusión del tejido. Este tipo de imagen radiológica resalta tejidos blandos como la materia gris (tonalidad gris), materia blanca (tonalidad ligeramente hipointensa respecto a la materia gris), tejido graso (poca intensidad debido a la poca presencia de agua), fluido cerebro espinal (tonalidad oscura) entre otros tejidos blandos.

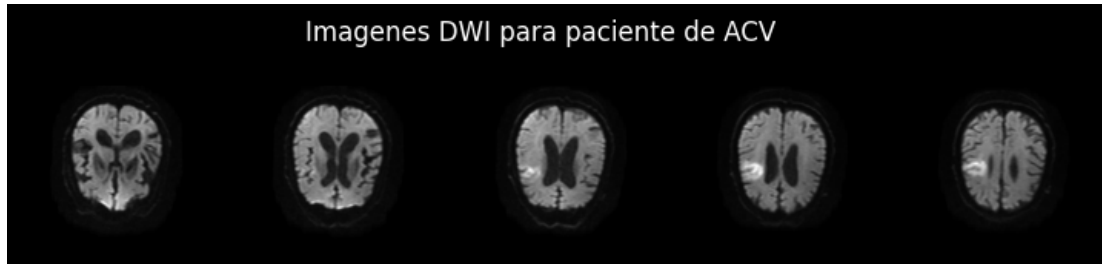
Para estas imágenes, el flujo restringido de agua (característico de ACV isquémico, cáncer y otras patologías) aparece como una señal hiperintensa. Sin embargo, una región hiperintensa no necesariamente está relacionada con un decremento en la perfusión de la zona.

- **Imágenes de coeficiente aparente de difusión (ADC por sus siglas en inglés):**

Las imágenes son mostradas como un mapa paramétrico que cuantifica la magnitud del movimiento de las moléculas de agua dentro de un tejido. Esta imagen se expresa en  $mm^2/s$  y un valor bajo en esta medición es un indicativo fuerte de perfusión restringida. Al estar directamente relacionada con el movimiento de las partículas de agua dentro del

## Figura 2

*Mapa de difusión isotrópico para paciente de ACV*

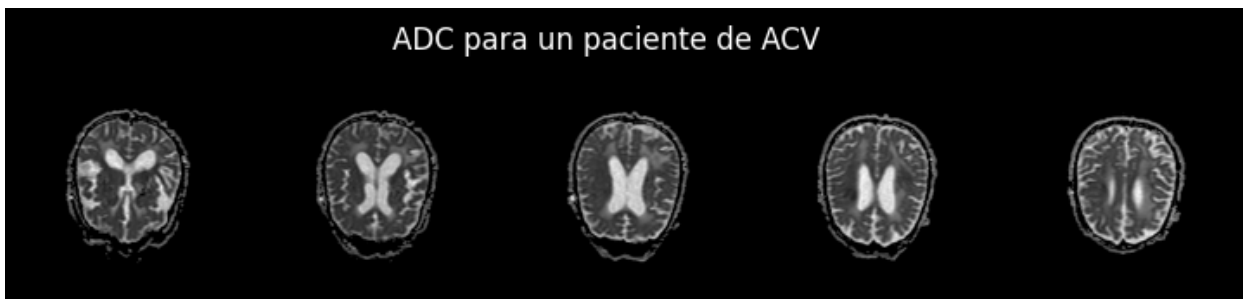


*Nota.* En este se pueden observar cortes para un paciente con ACV isquémico en imágenes de DWI. La lesión isquémica se puede observar como regiones de intensidad mayor respecto al resto del tejido cerebral.

cerebro, un valor de hipointenso en esta imagen es indicativo de un nivel bajo de perfusión y, por lo tanto, indicativo de ACV isquémico.

## Figura 3

*Coefficiente Aparente de Difusión para paciente de ACV*



*Nota.* En este se pueden observar cortes para un paciente con ACV isquémico en imágenes de ADC. La lesión isquémica se puede observar como regiones de intensidad menor respecto al resto del tejido cerebral.

Las imágenes multiparametricas se codifican de acuerdo a la respuesta del resonador, discretizando diferentes mapas de acuerdo al valor  $b$ , que hace referencia a la naturaleza de los campos magnéticos usados para su captura. Por ejemplo, un  $b = 0$  es un mapa paramétrico T2, que representa principalmente la estructura de las regiones del cerebro, mientras que un  $b = 1000$  es un mapa de difusión isotrópico de DWI. El ADC entonces surge del cálculo entre dos imágenes con diferente parámetro  $b$ . El análisis de lesiones de

ACV, por lo tanto, requiere el análisis integrado de mapas con diferentes parámetros de  $b$ , siendo los dos más comunes los mapas de DWI y ADC para determinar cambios en la hemodinámica cerebral dados en las primeras horas luego de un ACV. Las dos imágenes son necesarias para confirmar una lesión isquémica.

- **Imágenes de recuperación de inversión atenuada por fluidos (FLAIR por sus siglas en inglés):**

las imágenes FLAIR representan una herramienta esencial para la evaluación precisa y temprana de las lesiones cerebrales. Esta secuencia de resonancia magnética, al suprimir la señal del líquido cefalorraquídeo, resalta de manera óptima los cambios isquémicos en el tejido cerebral, permitiendo una detección sensible de las áreas afectadas. Además, las imágenes FLAIR pueden proporcionar información crucial para estimar el tiempo de evolución de la lesión isquémica, ya que ciertos hallazgos radiológicos como la hiperintensidad en el núcleo del infarto y la presencia de signos de penumbra isquémica pueden indicar la antigüedad de la lesión. Esta capacidad para dar un estimado del tiempo de la lesión es fundamental en la toma de decisiones clínicas, especialmente en el contexto de la terapia de reperfusión, donde el tiempo es un factor crítico para determinar la elegibilidad del paciente y el beneficio potencial de la intervención.

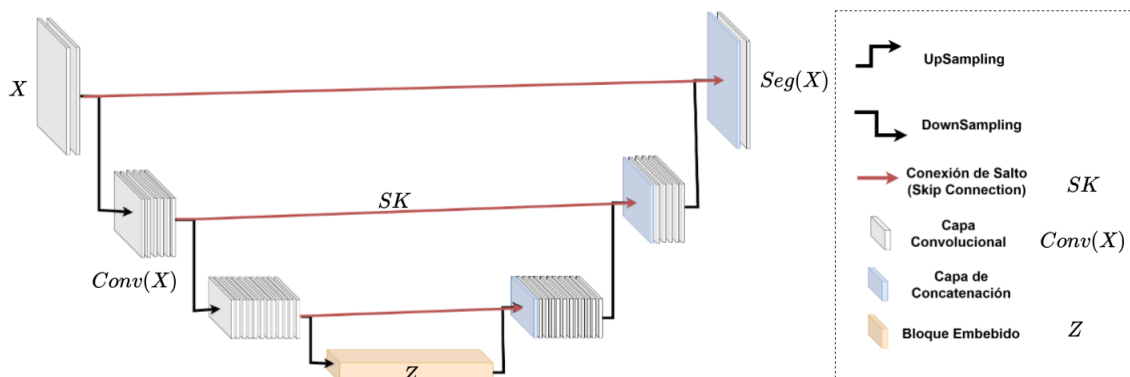
## 1.2. U-Net: una arquitectura para la segmentación de imágenes médicas

La segmentación de imágenes radiológicas en aprendizaje profundo se refiere al proceso de dividir una imagen en múltiples regiones llamadas segmentos, con el fin de identificar objetos específicos en la imagen. Por ejemplo, en el caso del ACV isquémico, el objetivo es demarcar la zona correspondiente al lugar de la lesión.

A lo largo de los años, los sistemas de diagnóstico asistido por computador han ido avanzando significativamente en la integración de tareas de segmentación en los flujos de trabajo clínicos. Uno de los enfoques más prometedores en este campo es la U-Net, una arquitectura

de aprendizaje profundo ampliamente utilizada para la segmentación de imágenes en diversos dominios médicos. La arquitectura U-Net sigue una estructura codificador-decodificador, siguiendo una arquitectura típica de contracción y otra de expansión de la representación. La arquitectura de contracción captura el contexto de la imagen utilizando capas convolucionales, mientras que la ruta de expansión produce una segmentación a partir de características de baja dimensión. Además, estas redes utilizan conexiones de salto entre el codificador y el decodificador para obtener segmentaciones más precisas. Las redes U-Net se basan en una representación convolucional que resume las características a partir de mapas de activación organizados jerárquicamente en cada capa. Estas redes especializadas tienen varias ventajas, como producir menos conexiones, compartir parámetros y una representación localmente equivariante que las hace localmente invariantes a las traslaciones, lo que les permite manejar entradas de longitudes variables. Sin embargo, el sesgo inductivo convolucional limita la capacidad de los modelos para explotar las relaciones no locales presentes en las imágenes. Para abordar este problema, muchos trabajos recientes han adoptado módulos de atención que pueden capturar patrones no locales necesarios para modelar formas complejas como las lesiones por ACV.

**Figura 4**  
Arquitectura U-Net



*Nota.* Se compone de un camino de contracción para capturar características profundas y un camino de expansión para generar segmentaciones detalladas. Sus conexiones residuales entre capas, conocidas como *Skip Connections*, permiten una mejor propagación de la información. La estructura en forma de U facilita la precisión en la segmentación de objetos en imágenes médicas y biomédicas.

### 1.3. Estrategias para segmentar lesiones de ACV isquémico

Por ejemplo, Shin *et al.* propuso un autocodificador con e-bloques (*efficient-blocks*) y módulos de atención para incluir patrones visuales de no locales de las modalidades MRI <sup>6</sup>. Por otro lado, Wang *et al.* propuso un esquema de redes generativas adversarias (GAN) para estimar lesiones de ACV isquémico, subagudo y crónico desde modalidades T2w. El esquema GAN está compuesto por un módulo de conexión de similitudes que captura características multiescala para mantener una consistencia perceptual, y así reducir la dependencia de datos completamente anotados <sup>7</sup>. Por ejemplo, Zhou *et al.* propusieron la D-UNet, una red neuronal de bajo costo computacional, que combina convoluciones 2D y 3D para extraer características texturales y espaciales de modalidades MRI <sup>16</sup>. Esta arquitectura también involucra una función de costo Enhance Mixed, que combina las pérdidas Focal y Dice. También, Qi *et al.* propusieron X-Net, una arquitectura codificador-decodificador que incluye bloques-X y operaciones no locales FSM (*Feature Similarity Map*) para segmentar lesiones de ACV isquémico desde imágenes T2w <sup>17</sup>. Esta arquitectura permite extraer información contextual densa para mejorar la estimación de las lesiones, mientras mantiene una representación compacta utilizando convoluciones separables en profundidad. Además, Zhang *et al.* uso un enfoque en dos etapas que consiste en utilizar una U-Net 3D para analizar parches volumétricos de tejido sano y con lesión, muestreados de forma

---

<sup>16</sup> Yongjin ZHOU et al. “D-UNET: a Dimension-Fusion U shape network for chronic stroke lesion segmentation”. In: *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 18.3 (May 2021), pp. 940–950. DOI: 10.1109/tcbb.2019.2939522.

<sup>17</sup> Kehan QI et al. *X-NET: Brain stroke lesion segmentation based on depthwise Separable convolution and Long-Range Dependencies*. Jan. 2019, pp. 247–255. DOI: 10.1007/978-3-030-32248-9\_28.

aleatoria<sup>18</sup>. Para la segunda etapa, las predicciones obtenidas se someten a un sub muestreo por ventana deslizante y se adicionan mediante bloques de atención a una U-Net 2D que proporciona la segmentación de la lesión de ACV isquémico. A pesar de obtener buenos resultados, estos trabajos llevan a cabo el análisis desde una única modalidad MRI, siendo más sensibles a un incremento de falsos positivos en las predicciones por no incluir información complementaria de otras modalidades.

En la literatura también se encuentran trabajos que construyen arquitecturas para la estimación de las lesiones de ACV isquémico, desde múltiples modalidades MRI que aportan información complementaria. Particularmente, existen trabajos que construyen autocodificadores que procesan múltiples modalidades de resonancia magnética a partir de una concatenación en la entrada. Por ejemplo, Kamnitsas *et al.*<sup>19</sup> implementó una red convolucional dual que explota la información contextual local y global de modalidades FLAIR y DWI, por medio de un procesamiento multi-escala. Otros enfoques han explorado la posibilidad de trabajar a partir de la concatenación de parches de modalidades multiparametricas MRI, en lugar de la modalidad médica completa<sup>20</sup>. De esta forma, se toman la misma cantidad de parches de tejido sano e infartado para llevar a cabo un entrenamiento más balanceado del modelo. Esta estrategia es útil para abordar el problema de desbalance de clases, sin embargo, impone un procesamiento local, perdiendo las observaciones geométricas globales, que pueden ser clave para recuperar adecuadamente la forma de las lesiones.

También encontramos estrategias que integran mecanismos para calcular relaciones glob-

---

<sup>18</sup> Yue ZHANG et al. *A 3D+2D CNN approach incorporating boundary loss for stroke lesion segmentation*. Jan. 2020, pp. 101–110. DOI: 10.1007/978-3-030-59861-7\_11.

<sup>19</sup> Konstantinos KAMNITSAS et al. “Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation”. In: *Medical Image Analysis* 36 (Feb. 2017), pp. 61–78. DOI: 10.1016/j.media.2016.10.004.

<sup>20</sup> Albert CLÈRIGUES et al. “Acute and sub-acute stroke lesion segmentation from multimodal MRI”. in: *Computer Methods and Programs in Biomedicine* 194 (2020). DOI: 10.1016/j.cmpb.2020.105521. arXiv: 1810.13304.

ales entre las características profundas calculadas a partir de la concatenación de modalidades de resonancia magnética. Por ejemplo, Liu *et al.* propuso una U-Net con mecanismos de atención residuales (DRANet) a partir de convolucionales dilatadas para segmentar y cuantificar de forma simultánea la lesión de ACV isquémico e hiperintensidades de materia blanca <sup>21</sup>. Así mismo, Ou *et al.* implementa capas Lambda+ para calcular coeficientes globales, locales e intercanal para segmentar lesiones isquémicas desde imágenes MRI ponderadas por difusión. Estas capas siguen el mismo principio que los mecanismos de atención pero con una complejidad de cálculo lineal <sup>22</sup>. Ashtar *et al.* logra aprovechar la complementariedad entre imágenes DWI, ADC y FLAIR mediante una arquitectura multimodal por medio de una capa NMF (*Non-negative Matrix Factorization*) y una operación de conversión entre tensores y matrices por desplazamiento de ventana (shifted window matricize operation)<sup>23</sup>. Bal *et al.*, implemento una red convolucional con 2 caminos con el objetivo de explotar las relaciones locales y no locales de imágenes de DWI y FLAIR. Este enfoque permite que el modelo tenga un entendimiento más detallado de las partes más detalladas de la lesión isquémica. A pesar de que estos trabajos estén diseñados explícitamente para explotar información multimodal sobre modalidades de resonancia magnética, la captura de patrones visuales asociados a lesiones de ACV isquémico mediante una simple concatenación puede verse limitada. De forma alternativa, Pinto *et al.* obtuvo características profundas que resumen información hemodinámica y del núcleo de la lesión mediante Máquinas de Boltzmann en dos ramas separadas para posteriormente entrenar una red convolucional recurrente. Esta red aprovecha la complementariedad entre los mapas obtenidos, resaltando los patrones visuales de las modalidades relacionadas y la complemen-

---

<sup>21</sup> Liangliang LIU et al. “Attention convolutional neural network for accurate segmentation and quantification of lesions in ischemic stroke disease”. In: *Medical Image Analysis* 65 (2020), p. 101791.

<sup>22</sup> Yanglan OU et al. *LambdaUNET: 2.5D stroke lesion segmentation of Diffusion-Weighted MR Images*. Jan. 2021, pp. 731–741. DOI: 10.1007/978-3-030-87193-2\_69.

<sup>23</sup> Pooya ASHTARI et al. “Factorizer: A scalable interpretable approach to context modeling for medical image segmentation”. In: *Medical image analysis* 84 (2023), p. 102706.

tariedad entre diferentes modalidades <sup>24</sup>. Por otro lado, Gómez *et al.* utilizó una red U-Net dual con bloques residuales y atención cruzada aditiva sobre las modalidades ADC y DWI para segmentar la lesión isquémica. A pesar de los notables avances en las estrategias de aprendizaje, así como la validación del aporte de la integración multimodal, aún el problema de segmentación y caracterización de la lesión permanece abierto. De hecho, aún se requieren herramientas con mayor capacidad de definición, que permitan una estimación apropiada del volumen de la lesión. También se deben crear representaciones que exploten hallazgos radiológicos desde diferentes modalidades, emulando análisis típicos desarrollados por expertos neuro-radiólogos.

---

<sup>24</sup> Adriano PINTO et al. “Combining unsupervised and supervised learning for predicting the final stroke lesion”. In: *Medical image analysis* 69 (2021), p. 101888.

## 2. PROBLEMA DE INVESTIGACIÓN

El ACV isquémico es una enfermedad cuyo diagnóstico prevalece como un reto a la salud mundial. El principal marcador diagnóstico del ACV es el volumen de la lesión, cuantificado como un cambio de perfusión en secuencias de resonancia magnética (RM). Sin embargo, este proceso de delineación suele ser lento (reportándose tiempos de aproximadamente 15 minutos por paciente), y además está sujeto a la experticia del profesional que realiza la tarea. Además, como se reporta en la literatura, existe acuerdo moderado entre delineaciones realizadas por expertos ( $\kappa = 65\%$ ). Es por ello, que métodos computacionales han emergido como una potencial alternativa para reducir el sesgo en las delineaciones, así como también mejorar los tiempos de segmentación de las mismas. Además, estos modelos típicamente son ajustados para realizar segmentaciones sobre modalidades únicas de observación, y limitándose a validaciones con un solo observador. Sin embargo, en el estándar clínico, incluso los expertos neuro-radiólogos observan hallazgos radiológicos sobre diferentes secuencias de RM (ADC, DWI, FLAIR) para definir los contornos de la lesión. Esta información es complementaria entre diferentes modalidades y puede ser clave para el modelamiento de la segmentación de lesiones.

### **Pregunta de Investigación**

¿Cómo integrar múltiples secuencias de resonancia magnética para aprovechar hallazgos radiológicos complementarios, que puedan incidir en la segmentación de lesiones de ACV isquémico?

### 3. OBJETIVOS

#### Objetivo General

- Implementar una arquitectura multimodal para la segmentación de lesiones de ACV isquémico sobre secuencias de resonancia magnética.

#### Objetivos Específicos

- Procesar un conjunto de datos con múltiples estudios de resonancia magnética que tengan anotaciones de lesiones de ACV isquémico.
- Implementar arquitecturas para la segmentación de lesiones de ACV isquémico, ajustadas desde una única modalidad de resonancia magnética.
- Implementar una arquitectura de aprendizaje profundo multimodal que aproveche la complementariedad entre distintas secuencias de resonancia magnética.
- Validar las arquitecturas propuestas con respecto a la capacidad de segmentar lesiones de ACV isquémico utilizando anotaciones realizadas por expertos radiólogos.

## 4. MÉTODO PROPUESTO

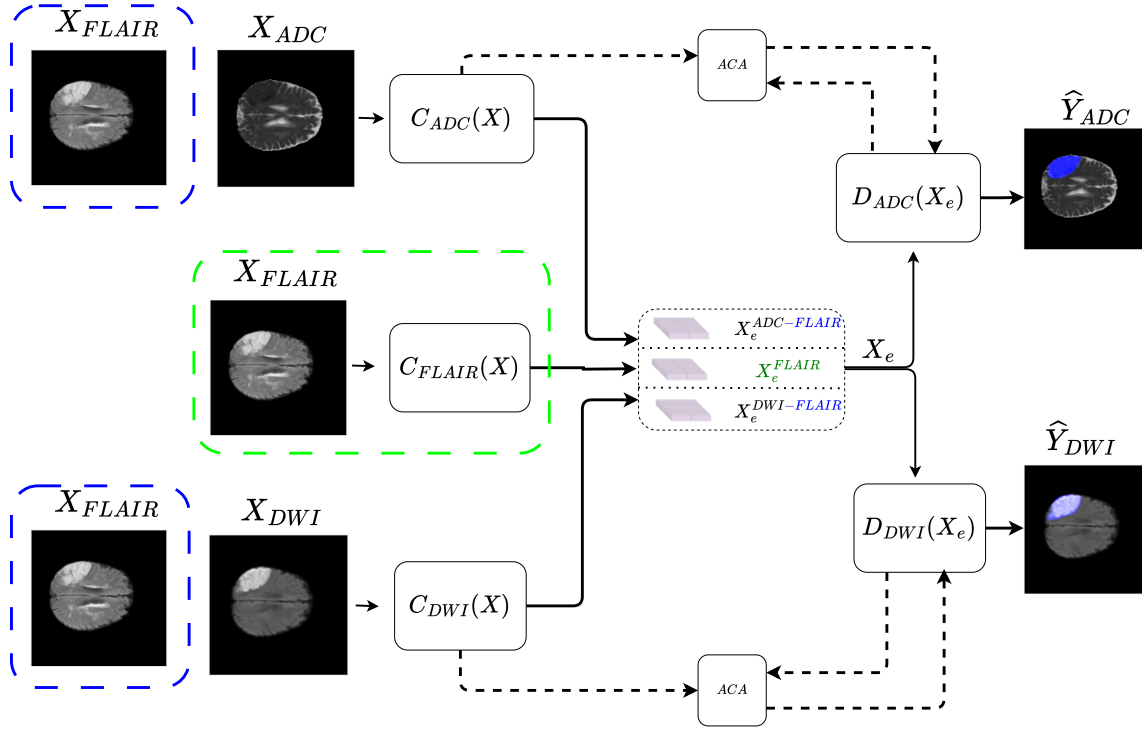
En este trabajo se desarrollaron y exploraron representaciones unimodales y multimodales de múltiples ramas para ADC y DWI, que pueden estar complementadas con mapas FLAIR. La arquitectura de aprendizaje incluye una representación codificador-decodificador de múltiples caminos con capacidad para recuperar segmentaciones de ACV a partir de las modalidades de ADC y DWI. Ramas independientes son codificadas para cada modalidad y se fusionan en su representación embebida. A continuación, se desarrollan rutas de decodificación independientes para recuperar las lesiones del ACV, con el apoyo de módulos de atención cruzada de los codificadores respectivos. Además, en la arquitectura desarrollada, se implementó una tarea de segmentación múltiple (en ADC y DWI) para actualizar la representación profunda propuesta. La descripción general del enfoque propuesto, así como sus variantes para complementar con los mapas FLAIR se ilustra en la figura 5

### 4.1. Modelo multimodal y multisegmentación

Hoy en día, los autocodificadores son la arquitectura más utilizada para tratar la segmentación de lesiones de ACV. En concreto, la U-Net se compone de un codificador, en donde mediante filtros convolucionales, operaciones de normalización y reducción de dimensionalidad se extraen las características profundas. Así, se espera que la arquitectura aprenda a resaltar los patrones visuales relevantes de la lesión isquémica, codificándolos en un vector de características conocido como vector embebido. Desde esta proyección embebida de baja dimensión se hace la estimación y reconstrucción de la máscara de segmentación en el decodificador mediante operaciones de muestreo, filtros convolucionales y normalizaciones. Adicionalmente, a esto, la arquitectura U-Net incluye conexiones de salto para integrar características profundas, desde el codificador al decodificador, en el mismo nivel de procesamiento. En estos casos, la arquitectura ha demostrado ser ideal en problemas que requieren recuperar información estructural y

**Figura 5**

Representación gráfica de los métodos de integración multimodal para la segmentación de la lesión isquémica.



*Nota.* En azul se presenta la integración de imágenes FLAIR como modalidad auxiliar en la detección del ACV isquémico. En verde se presenta la integración tardía del FLAIR en el espacio latente mediante la extracción de características en un codificador dedicado. Adicionalmente, cada arquitectura consta de tres partes principales: el codificador ( $C(X)$ ), que caracteriza la información más relevante de la imagen de entrada y la transforma en un embebido de baja dimensión; el decodificador ( $D(X)$ ), que recupera la información de alta dimensión capturada en la representación embebida; y el módulo de atención  $ACA$ , que filtra y relaciona las características más importantes en ambas modalidades.

de forma global.

Este trabajo se usó como base de conocimiento esta arquitectura U-Net, proponiendo así adaptaciones multimodales que permitieran procesar múltiples ramas de procesamiento y robusteciendo su representación con módulos de atención. Además, la arquitectura propuesta incluye ramas independientes de estimación y reconstrucción de la lesión, logrando ser efectivas en tareas de múltiples segmentaciones, provenientes de delineaciones desde diferentes observaciones

(típicamente ADC y DWI). Particularmente, en el modelo base se consideran el procesamiento de los estudios DWI y MRI, los cuales son procesados de forma independiente, y cuyas observaciones cuentan con máscaras independientes, de acuerdo a los hallazgos radiológicos observados en cada secuencia imagenológica. Cabe recordar, que en un proceso típico de análisis por parte de un experto, estas imágenes son observadas en paralelo por el experto y la complementariedad de hallazgos en las dos imágenes define los contornos de la lesión isquémica.

Entonces, en este trabajo se implementaron codificadores independientes por cada rama de procesamiento, los cuales incluyen filtros convolucionales, operaciones de normalización y reducción de dimensionalidad, por cada modalidad, conduciendo a una representación embebida independiente. Considerando igual importancia entre las imágenes de entrada para establecer una lesión isquémica, las ramas de procesamiento son simétrica y proyectan la representación en igualdad de condiciones al espacio embebido. Los dos vectores embebidos (ADC y DWI) son concatenados para obtener un único vector que represente la lesión isquémica. Desde este vector fusionado se desarrollan ramas independientes de decodificación de la información. Estas ramas conducen a una reconstrucción de la máscara respectiva que contiene los hallazgos observados en esa modalidad.

Para robustecer el proceso de decodificación, en este trabajo se incluyó un modelo de atención aditiva cruzada en las conexiones de salto que pondera la contribución del codificador  $X_e^i$  a la representación del decodificador  $X_d^i$ , en cada nivel convolucional  $i$ . Las características refinadas del codificador se calculan como:

$$X_{re}^i = \sigma_2(W_{re}^T \sigma_1(W_e^T X_e^i + W_d^T X_d^i)) \cdot X_e^i$$

Este módulo de atención es utilizado en cada rama de procesamiento de forma independiente, permitiendo resaltar la información predominante desde el codificador, para apoyar la reconstrucción embebida en cada nivel de procesamiento. Así, la reconstrucción en cada modalidad independiente podrá generar segmentaciones asociadas a las observaciones en cada parámetro de MRI. Para dicho fin, la arquitectura implementada fue ajustada a través de una

función de pérdida multisegmentación para capturar patrones texturales complementarios de ambos parámetros de RM y mejorar la segmentación de la lesión. Esta función induce el aprendizaje de características altamente correlacionadas presentes en las imágenes ADC y DWI. En concreto, la función de pérdida se definió como la suma de las pérdidas de cada parámetro de RM de la siguiente manera:

$$\mathcal{L} = \mathcal{L}_{ADC} + \mathcal{L}_{DWI} \quad (1)$$

$$\mathcal{L}_{ADC} = \mathcal{L}_{CE}(\mathbf{Y}_{ADC}, \hat{\mathbf{Y}}_{ADC}) + \mathcal{L}_{DICE}(\mathbf{Y}_{ADC}, \hat{\mathbf{Y}}_{ADC})$$

$$\mathcal{L}_{DWI} = \mathcal{L}_{CE}(\mathbf{Y}_{DWI}, \hat{\mathbf{Y}}_{DWI}) + \mathcal{L}_{DICE}(\mathbf{Y}_{DWI}, \hat{\mathbf{Y}}_{DWI})$$

, donde  $\mathcal{L}_{CE}(\cdot, \cdot)$  y  $\mathcal{L}_{DICE}(\cdot, \cdot)$  son funciones de entropía cruzada y Dice calculadas de la siguiente manera:

$$\mathcal{L}_{CE}(\mathbf{Y}, \hat{\mathbf{Y}}, \mathbf{C}) = \mathbf{Y} \log(1 - \hat{\mathbf{Y}}) \mathbf{C} + (1 - \mathbf{Y}) \log(\hat{\mathbf{Y}}) \mathbf{C} \quad (2)$$

$$\mathcal{L}_{DICE}(\mathbf{Y}, \hat{\mathbf{Y}}) = 1 - \frac{2|\mathbf{Y} \cap \hat{\mathbf{Y}}|}{|\mathbf{Y}| + |\hat{\mathbf{Y}}|} \quad (3)$$

Aquí,  $\hat{\mathbf{Y}}_{ADC}$  y  $\hat{\mathbf{Y}}_{DWI}$  son las anotaciones para ADC y DWI, respectivamente. Cabe señalar que la segmentación del ACV isquémico está muy desequilibrada, lo que favorece la reconstrucción del fondo. Para superar este problema, cada pérdida de modalidad se pondera mediante un conjunto de mapas de peso de clase ( $\mathbf{C} \in \mathbb{N}^{H \times W}$ ) que aumentan la contribución de los vóxeles de lesión y contrarrestan el impacto negativo del desequilibrio de clase. A partir de esta función de pérdida de segmentación múltiple, el modelo puede capturar con mayor eficacia la información complementaria de ADC y DWI, lo que conduce a una mejor segmentación de la lesión.

A partir de esta configuración base multimodal, enfocada en la caracterización y delimitación de lesiones en ADC y DWI, se exploraron dos adaptaciones alternativas para incluir y medir la contribución de la modalidad de FLAIR. Como en el análisis radiológico realizado por

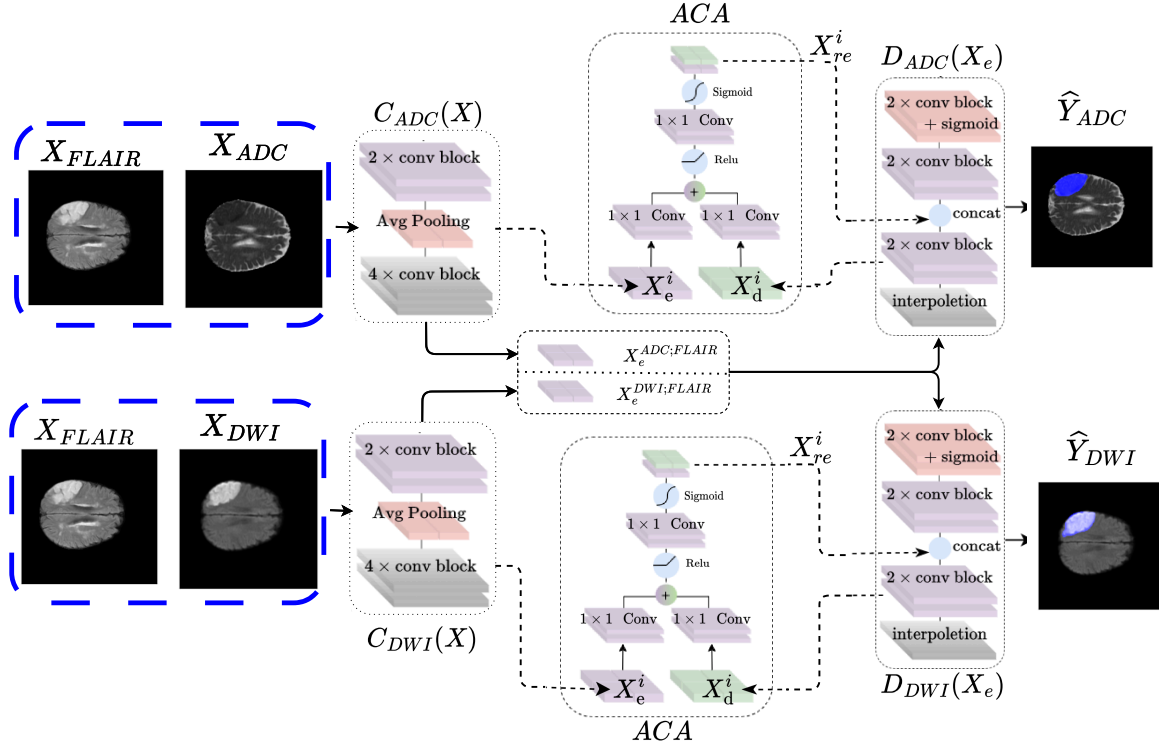
expertos, esta modalidad es una alternativa utilizada en algunos pacientes como mecanismo de tamizaje y comprobación de un hallazgo radiológico, asociado a una lesión isquémica. El primer enfoque usa una integración temprana de FLAIR, incluyéndolo como un canal adicional en la entrada del codificador para ambas imágenes (ADC y DWI). Por otro lado, se exploró una fusión tardía, donde se extraen las características profundas al espacio latente para concatenar todos los vectores embebidos. En las siguientes secciones se detalla el respectivo funcionamiento de cada uno de estos enfoques.

## **4.2. Fusión temprana del FLAIR**

Una primera alternativa a partir de la arquitectura implementada, fue la de incluir la secuencia FLAIR concatenada con las secuencias ADC y DWI de manera temprana. En este caso, el FLAIR se comporta como una modalidad auxiliar, fusionándola de manera temprana desde la entrada del codificador. La hipótesis detrás de esta fusión temprana es que el FLAIR funciona como un mapa de atención para lesiones mayores a seis horas, resaltando la lesión isquémica de forma similar al DWI, sin embargo, para lesiones menores a este periodo de tiempo al no presentar cambios significativos en intensidad llega a funcionar como un regularizador al momento de obtener las características profundas de la lesión isquémica. En la figura 6 se ilustra la adaptación de la arquitectura base para incluir las secuencias FLAIR de forma temprana.

Particularmente en esta fusión, la imagen de FLAIR puede ser adoptada en la entrada tanto del DWI, como del ADC, pero también se exploraron alternativas únicamente soportando un canal de información. En este caso, el mapa del FAIR entra como un canal adicional en la entrada. Luego las capas convolucionales operan sobre cada uno de estos canales, realizando una fusión local conjunta con el nuevo mapa. Así, por ejemplo, en la fusión con el DWI se emula el análisis conjunto que permite descartar una lesión isquémica temprana. En cada rama de procesamiento que se incluye el FLAIR, los patrones convolucionales pueden ser propagados hasta tener una representación en el vector embebido.

**Figura 6**  
*Representación gráfica de la fusión temprana del FLAIR.*



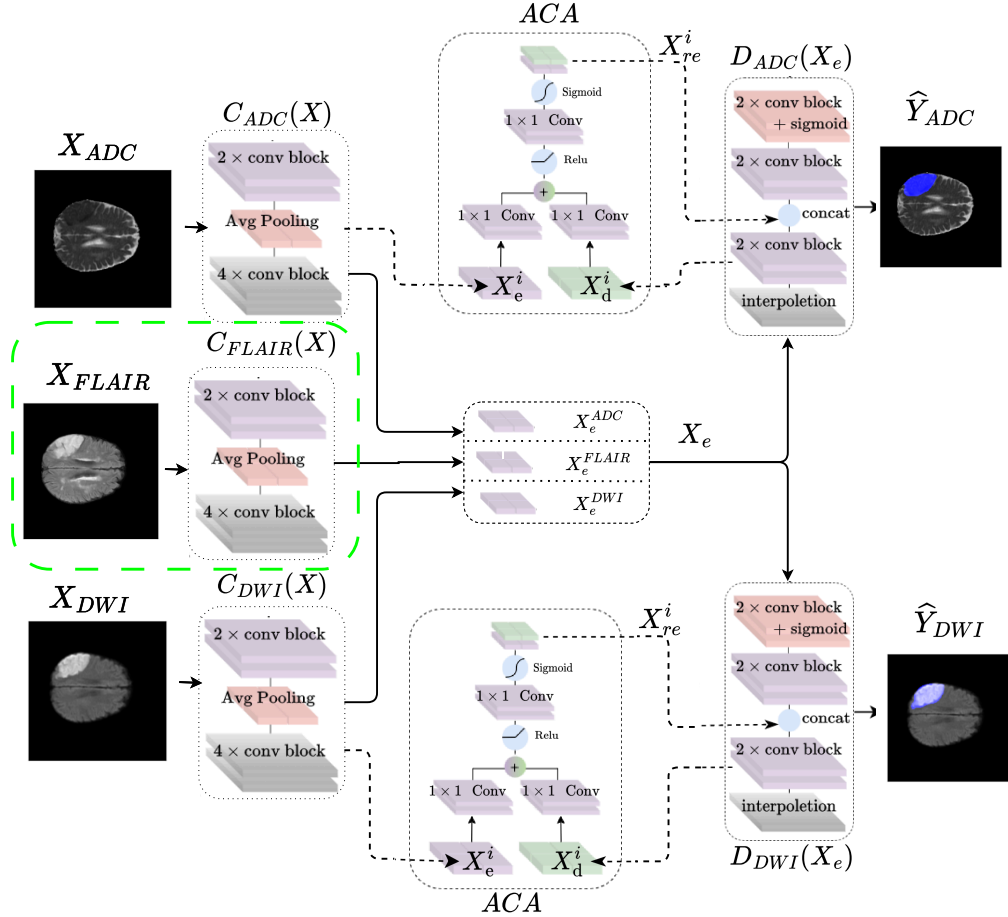
*Nota.* Este enfoque de integración del FLAIR busca concatenar esta modalidad a la entrada de los codificadores de ADC y DWI para extraer las características profundas de forma conjunta entre las modalidades.

### 4.3. Fusión tardía del FLAIR

Como alternativa a la fusión temprana de la modalidad del FLAIR, en la figura 7 se ilustra una alternativa propuesta en este trabajo, en la cual las secuencias de FLAIR se procesan y representan como una rama independiente de codificación, que proyecta su representación en un descriptor embebido. Específicamente, La fusión tardía extrae los patrones visuales del FLAIR como modalidad completamente independiente en un codificador convolucional, que mantiene la misma estructura de la modalidad ADC y DWI. Una vez cada rama independiente hace su proyección hacia un vector embebido, entonces los vectores se concatenan, formando una única representación de las lesiones isquémicas.

**Figura 7**

*Representación gráfica de la fusión tardía del FLAIR.*



*Nota.* El recuadro punteado verde muestra la variación principal de este enfoque, un codificador dedicado para la extracción de características profundas del FLAIR que posteriormente se fusionan en el espacio embebido.

En esta configuración, La rama de FLAIR no tiene un decodificador asociado, considerando que en la rutina clínica esta secuencia no es usada para medir o definir volumen en las lesiones isquémicas. Por lo tanto, en el conjunto de datos considerado no se cuentan con delineaciones sobre esta secuencia. Sin embargo, la propagación de las ramas de DWI y ADC impactan el ajuste del codificador FLAIR, permitiendo así ajustar la representación a maximizar el descriptor que asocia los hallazgos de la lesión isquémica.

## 5. CONFIGURACIÓN EXPERIMENTAL

### 5.1. Datos

Se realizó un estudio retrospectivo para validar la capacidad del enfoque propuesto, estimando la segmentación de lesiones de ACV isquémico. El estudio recogió 180 estudios de pacientes con síntomas de ACV en dos centros clínicos entre octubre de 2021 y septiembre de 2022. Cada estudio contiene secuencias ADC, DWI Y FLAIR con resoluciones que varían en cada eje como  $x,y=[0,5-0,9]$  y  $z=[6,0-7,2]$ mm para el ADC y DWI y  $x,y=[0,4-0,5]$  y  $z=[6,0-12,0]$ mm para FLAIR. Sólo se consideraron las imágenes DWI B1000. Los estudios se capturaron utilizando dos escáneres de RM, a saber, Toshiba Vantage Titan MRI y General Electric SIGNA Explorer, ambos con un campo magnético de 1,5T. Para cumplir los criterios de inclusión, los pacientes debían ser mayores de 18 años y no mostrar signos de hemorragia cerebral. Los pacientes que mostraban indicios de reperfusión parcial no fueron excluidos del estudio. Tres radiólogos con más de cinco años de experiencia analizaron las imágenes diagnósticas y anotaron manualmente las lesiones de ACV. Las anotaciones se realizaron de forma independiente utilizando el software MRICroGL<sup>25</sup> y utilizando la historia clínica al ingreso. Tras revisar los estudios de imagen, los estudios se clasificaron en estudios de control (n=17) o ACV isquémico (n=163). Se incluyeron pacientes de control con síntomas de ACV para diversificar las muestras de tejido, mejorando potencialmente la capacidad de nuestros modelos para segmentar el tejido de la lesión de ACV.

Para entrenar y evaluar la capacidad de captar información multimodal de nuestro enfoque propuesto, el conjunto de datos se estratificó en dos subconjuntos para entrenamiento (n=126) y prueba (n=54). Para controlar la capacidad del modelo para segmentar lesiones de ACV isquémico en datos no vistos y establecer un criterio para la selección del modelo, se

---

<sup>25</sup> Xiangrui et al. LI. “The first step for neuroimaging data analysis: DICOM to NIfTI conversion”. In: *Journal of neuroscience methods* 264 (2016), pp. 47–56.

creó un subconjunto de validación a partir de 27 estudios de la partición de entrenamiento. El conjunto de prueba se diseñó para contener anotaciones manuales de los tres expertos, lo que permitió cuantificar las diferencias entre las anotaciones de las lesiones. En cuanto al preprocesamiento, removi6 el cráneo a todas las secuencias, se realizó un registro r6gido para alinear la secuencia FLAIR con la ADC y, por 6ltimo, se remuestrearon todas las imágenes a la resoluci6n media de v6xel del conjunto de datos de entrenamiento ( $x,y,z=[0,9, 0,9, 6,0]$ ).

## 5.2. Configuraci6n de la arquitectura

En cuanto al modelo, la arquitectura propuesta consta de cinco niveles convolucionales (32, 64, 128 y 256 filtros para ambas redes y 512 filtros en el cuello de botella) con bloques residuales formados por dos conjuntos de  $3 \times 3$  de convoluci6n estriada, normalizaci6n de grupos (grupos=8), y activaci6n SiLU. Para el entrenamiento, la arquitectura se actualiz6 a lo largo de 200 6pocas, siguiendo la p6rdida de multisegmentaci6n propuesta, un optimizador AdamW con  $\gamma = 2$  y  $\alpha = 0,25$  y un tama6o de lote de 32. Los valores del mapa de pesos de clase se fijaron en 0,5 para los v6xeles no lesionales y en 3 para los v6xeles lesionales. La tasa de aprendizaje inicial se fij6 en  $5e-4$  y el decaimiento de los pesos en  $3e-5$ . Se aplic6 una reducci6n polin6mica a la tasa de aprendizaje a trav6s de las 6pocas. Se implement6 una reducci6n polin6mica para estabilizar el proceso de entrenamiento, ajustando la tasa de aprendizaje a medida que progresaba el entrenamiento. Los estudios completos se normalizaron y redimensionaron a  $128 \times 128$  p6xeles.

## 5.3. M6tricas de evaluaci6n

Para evaluar la calidad de las estimaciones de las lesiones isqu6micas, se tuvieron en cuenta m6tricas de solapamiento, cl6nicas y de l6mites. A continuaci6n se detallan las m6tricas utilizadas:

- **Diferencia absoluta de volumen (AVD, por sus siglas en ingl6s):** Refleja la exactitud con la que el modelo predice el volumen de la lesi6n, un biomarcador clave en la

definición del tratamiento y pronóstico del paciente.

$$AVD = |(N_{real} \times V_{voxel}) - (N_{pred} \times V_{voxel})|$$

Donde  $N_{real}$  es el número total de vóxeles (unidad cúbica que compone un objeto tridimensional) clasificados como lesión en las marcaciones del experto.  $N_{pred}$  es el número total de vóxeles clasificados como lesión por el modelo de predicción.  $V_{voxel}$  es el volumen de un solo vóxel (por ejemplo, si la resolución de la imagen es 1 mm en cada dirección, entonces  $V_{voxel} = 1mm^3$

- **Diferencia absoluta de lesión (ALD, por sus siglas en inglés):** Refleja la capacidad del modelo para cuantificar con precisión la cantidad de lesiones de ACV en un estudio imagenológico.

$$ALD = |L_{real} - L_{pred}|$$

Donde  $L_{real}$  es el número de lesiones identificadas por el experto,  $L_{pred}$  es el número de lesiones identificadas por el modelo. Se calculó  $L_{real}$  y  $L_{pred}$  mediante un análisis de componentes conectados. Este análisis etiqueta cada grupo conectado de vóxeles identificados como lesión, como un componente individual, representando así una lesión única. Tras etiquetar estos componentes, contamos el número total de componentes únicos en cada máscara binaria.

- **F1-score de detección de lesión (LF1):** Es una métrica que combina la precisión y la sensibilidad para medir la capacidad de detección de lesiones. Un F1-score alto indica que el modelo es efectivo en la detección precisa de lesiones tanto en términos de localización como de cantidad.

$$LF1 = \frac{2TP}{2TP + FP + FN}$$

Al igual que en la métrica ALD, se llevó a cabo un análisis de componentes conectados para determinar las lesiones presentes en cada una de las marcaciones. En este caso, las cardinalidades básicas son definidas como  $TP$  el número de lesiones de la marcación del experto que se solapan al menos en un vóxel con las lesiones de la estimación del modelo,  $FP$  el número de lesiones de la estimación del modelo donde ningún vóxel se solapa con lesiones de la marcación del experto,  $FN$  el número de lesiones de la marcación del experto que no tiene ningún vóxel que se solape con lesiones de la estimación del modelo.

- **Distancia de Hausdorff (HD, por sus siglas en inglés):** Esta métrica mide la mayor de todas las distancias de un conjunto de puntos a otro. En el contexto de la segmentación de imágenes médicas, evalúa la máxima distancia entre los bordes de las lesiones segmentadas por el modelo y los bordes de las lesiones anotados por un radiólogo. Proporciona una estimación de la discrepancia espacial más significativa entre las dos segmentaciones, destacando la precisión con la que el modelo puede replicar la forma y el contorno de las lesiones reales.

$$HD = \max\left\{\max_{a \in A_s} \min_{b \in B_s} d(a, b), \max_{b \in B_s} \min_{a \in A_s} d(a, b)\right\}$$

Donde  $A_s$  es el conjunto de todos los puntos en los bordes de las lesiones reales ( $Y$ ),  $B_s$  es el conjunto de todos los puntos en los bordes de las lesiones predichas por el modelo ( $\hat{Y}$ ) y  $d(a, b)$  es la distancia euclidiana entre los puntos  $a$  y  $b$ . Esta métrica es especialmente útil para identificar casos donde la predicción puede estar cerca en términos de volumen o área general, con respecto a las marcaciones de expertos, pero difiere significativamente en términos de la localización exacta de los bordes de la lesión. Un valor más bajo de HD indica una mayor similitud entre las formas segmentadas, mientras que un valor más alto indica una discrepancia mayor.

- **Dice Score (DSC):** También llamado índice de solapamiento, es la métrica más utilizada para validar las segmentaciones de volumen médico. Además de la comparación

directa entre las segmentaciones automáticas y las hechas por expertos. Este coeficiente es matemáticamente equivalente al *F1-Score*.

$$DICE = \frac{2TP}{2TP + FP + FN}$$

#### 5.4. Evaluación del desempeño en la segmentación de lesiones de ACV isquémico

Para cuantificar el aporte de cada modalidad en el conjunto de datos y validar la capacidad de las estrategias implementadas para segmentar lesiones de ACV isquémico, se utilizó el conjunto de pruebas anotado por tres expertos. Inicialmente, se entrenó cada modelo de inteligencia artificial con una configuración específica de entradas y anotaciones, y se calculó el Dice score medio al final de cada *epoch* en la partición de validación para monitorear el rendimiento. Si el Dice score mejoraba, se guardaban los pesos del modelo en ese momento, asegurando que al final del entrenamiento se conservaran los que maximizaran el solapamiento en datos no vistos anteriormente. Durante la fase de inferencia, se cargaron para cada paciente la configuración de entradas y los mejores pesos del modelo para estimar las lesiones de ACV, comparando los resultados con las anotaciones de los expertos y calculando la media y desviación estándar de las métricas obtenidas. Este proceso proporcionó una evaluación detallada de la precisión y consistencia del modelo en la segmentación de lesiones isquémicas.

Durante la validación se utilizó una partición a partir del conjunto de datos destinado para el entrenamiento del modelo, en total 27 pacientes se excluyeron con el objetivo de validar el desempeño en cada etapa del aprendizaje. Los resultados reportados se obtuvieron a partir de las predicciones provenientes de los modelos con mejor promedio de *dice score* para las marcaciones de ADC y DWI sobre este conjunto de validación.

Los experimentos elegidos para medir la contribución del FLAIR a la tarea de segmentación se dividieron entre los dos enfoques tomados. Para la fusión temprana, se iteró sobre las diferentes opciones de concatenación para cada uno de los codificadores (ramas), dando como resultado las siguientes configuraciones para los modelos probados.

- Rama 1: ADC - Rama 2: DWI
- Rama 1: ADC y DWI - Rama 2: FLAIR
- Rama 1: ADC y FLAIR - Rama 2: DWI y FLAIR
- Rama 1: ADC - Rama 2: FLAIR y DWI

## 6. EVALUACIÓN Y RESULTADOS

Con el objetivo de medir la complementariedad entre las modalidades de ADC, DWI y FLAIR, en este trabajo se exploraron diferentes configuraciones de las arquitecturas multimodales desarrolladas, considerando además la contribución de la secuencia FLAIR. Además, se consideró la evaluación de una arquitectura unimodal para reportar el desempeño de representación por cada modalidad independiente. La inclusión de FLAIR como modalidad auxiliar se planteó como fusión temprana en la combinación de las diferentes entradas posibles. También se planteó una arquitectura multimodal con tres codificadores, realizando la extracción de características profundas de cada imagen de forma independiente. En las siguientes subsecciones, se reportan los resultados obtenidos de los modelos desarrollados y líneas bases, para distintas configuraciones de secuencias de entrada, con respecto a los hallazgos radiológicos anotados sobre ADC y DWI.

### 6.1. Evaluación de arquitecturas unimodales

Con el fin de cuantificar el aporte de cada modalidad presente en el conjunto de datos, y su correlación con las anotaciones de expertos realizadas sobre ADC y DWI, se llevaron a cabo múltiples experimentos a partir de configuraciones unimodales. Más específicamente, se entrenó la arquitectura U-Net que incluye los módulos de atención propuestos en las conexiones de salto, siguiendo las anotaciones de expertos en una única modalidad. En la Tabla 1 se muestran las métricas Dice score y HD calculadas a partir de las anotaciones de cada uno de los expertos en ambas modalidades.

La Tabla 1 muestra que la modalidad DWI alcanzó los mejores resultados en las métricas de Dice score y HD para todos los radiólogos. Como se esperaba, FLAIR mostró los menores índices de superposición y las mayores distancias en todas las evaluaciones; no obstante, se evidenció una mayor correlación con los hallazgos en DWI que con los de ADC. En esta misma

**Tabla 1***Métricas computacionales para modelos U-Net unimodales*

Mascara de Evaluacion	Entradas	R1		R2		R3	
		Dice $\uparrow$	HD $\downarrow$	Dice $\uparrow$	HD $\downarrow$	Dice $\uparrow$	HD $\downarrow$
ADC	ADC	<u>0.37<math>\pm</math>0.29</u>	<u>44.81<math>\pm</math>39.33</u>	<u>0.39<math>\pm</math>0.3</u>	<b>41.64<math>\pm</math>37.93</b>	<u>0.36<math>\pm</math>0.28</u>	<u>48.74<math>\pm</math>39.71</u>
	DWI	<b>0.52<math>\pm</math>0.29</b>	<b>35.3<math>\pm</math>34.26</b>	<b>0.5<math>\pm</math>0.28</b>	<u>45.02<math>\pm</math>38.63</u>	<b>0.46<math>\pm</math>0.25</b>	<b>47.95<math>\pm</math>39.31</b>
	FLAIR	0.04 $\pm$ 0.19	88.14 $\pm$ 28.15	0.06 $\pm$ 0.23	88.9 $\pm$ 35.74	0.04 $\pm$ 0.19	85.64 $\pm$ 26.06
DWI	ADC	<u>0.37<math>\pm</math>0.26</u>	<u>49.81<math>\pm</math>31.02</u>	<u>0.39<math>\pm</math>0.29</u>	<u>51.35<math>\pm</math>30.95</u>	<u>0.4<math>\pm</math>0.26</u>	<u>64.68<math>\pm</math>39.15</u>
	DWI	<b>0.57<math>\pm</math>0.28</b>	<b>36.61<math>\pm</math>28.87</b>	<b>0.54<math>\pm</math>0.29</b>	<b>46.7<math>\pm</math>36.88</b>	<b>0.53<math>\pm</math>0.28</b>	<b>56.46<math>\pm</math>41.62</b>
	FLAIR	0.14 $\pm$ 0.24	70.77 $\pm$ 31.55	0.14 $\pm$ 0.24	77.91 $\pm$ 32.88	0.13 $\pm$ 0.24	79.69 $\pm$ 34.08

*Nota.* Los mejores resultados según las anotaciones de ADC y DWI de cada experto están destacados en negrita, y los segundos mejores, subrayados.

línea, cabe destacar que las modalidades ADC y DWI de entrada mostraron un desempeño similar cuando se optimizaron usando tanto las máscaras delineadas para ADC como para DWI. Específicamente, los modelos que utilizaron ADC como entrada demostraron una habilidad similar para detectar lesiones, independientemente de si se optimizaron con las anotaciones propias de ADC o con las de DWI (diferencia entre radiólogos promedio en Dice=0.013 y HD=10.33). Por otro lado, los modelos con DWI en la entrada presentaron variaciones más significativas en su rendimiento cuando la optimización se realizó con base en amas anotaciones (diferencia promedio en Dice=0.053 y HD=3.83). Es de resaltar la consistencia en las métricas calculadas para diferentes expertos.

Para mejorar la evaluación del impacto de las modalidades de imagen en un conjunto de datos relacionados con ADC y DWI, se incorporaron métricas de interés clínico en los experimentos. Este enfoque permite una valoración más integral de la precisión diagnóstica y la utilidad clínica de cada modalidad, facilitando la identificación de sus fortalezas y limitaciones. En la Tabla 2 se muestran las métricas AVD, ALD y LF1 calculadas a partir de las anotaciones de cada uno de los expertos en ambas modalidades.

La Tabla 2 revela que los modelos entrenados a partir de entradas DWI no alcanzan el mismo nivel de dominio en las métricas AVD, ALD y LF1 como se evidenció en la Tabla 1 para los Dice score y HD. Los modelos con entradas de ADC o DWI, optimizados para cada modalidad específica, alcanzaron los resultados más altos en la métrica de ALD. Respecto a la métrica AVD,

**Tabla 2***Métricas médicas para modelos U-Net unimodales*

Mascara de Evaluacion	Entradas	R1			R2			R3		
		AVD ↓	ALD ↓	LF1 ↑	AVD ↓	ALD ↓	LF1 ↑	AVD ↓	ALD ↓	LF1 ↑
ADC	ADC	<b>6.08±7.96</b>	<b>1.69±1.92</b>	<u>0.53±0.35</u>	<b>6.43±8.52</b>	<b>1.59±1.88</b>	<u>0.54±0.35</u>	<u>7.91±10.3</u>	<b>2.65±2.52</b>	<u>0.5±0.32</u>
	DWI	<u>7.4±11.49</u>	2.19±3.54	<b>0.6±0.33</b>	<u>6.8±10.24</u>	2.65±4.15	<b>0.56±0.31</b>	<b>7.5±10.17</b>	<u>3.11±4.1</u>	<b>0.54±0.27</b>
	FLAIR	17.74±31.34	<u>2.15±1.99</u>	0.07±0.24	18.18±29.79	<u>2.06±2.35</u>	0.09±0.27	17.7±28.71	3.85±3.73	0.06±0.23
DWI	ADC	<u>9.01±14.16</u>	2.7±2.86	<u>0.41±0.29</u>	<u>9.96±14.02</u>	2.74±2.25	<u>0.4±0.28</u>	<u>9.03±12.3</u>	<u>2.74±2.42</u>	<u>0.38±0.26</u>
	DWI	<b>6.1±8.31</b>	<b>2.02±2.55</b>	<b>0.58±0.32</b>	<b>5.6±7.53</b>	<b>1.83±2.08</b>	<b>0.54±0.34</b>	<b>5.31±7.23</b>	<b>2.5±3.06</b>	<b>0.5±0.3</b>
	FLAIR	21.42±33.47	<u>2.56±3.53</u>	0.24±0.33	22.03±30.8	<u>2.56±3.46</u>	0.27±0.33	21.98±32.49	3.48±4.08	0.22±0.28

*Nota.* Los mejores resultados según las anotaciones de ADC y DWI de cada experto están destacados en negrita, y los segundos mejores, subrayados.

estos modelos también mostraron un desempeño superior, con la excepción de las evaluaciones basadas en las anotaciones de ADC de R3. En este caso, el modelo que utilizó exclusivamente DWI como entrada superó a los demás. En la métrica LF1, que evalúa la detección de múltiples segmentos, los modelos con entradas DWI obtuvieron los mejores resultados para cada métrica. Por su parte, FLAIR, a pesar de ser generalmente menos efectivo, obtuvo los segundos mejores resultados en ALD para R1 y R2, resaltando su eficacia en contextos específicos. Además, es destacable la consistencia observada en los resultados para R1 y R2, a pesar de no utilizar las máscaras de R2 en el entrenamiento. De manera opuesta, los resultados muestran una menor correlación entre las lesiones estimadas por los modelos con respecto a las anotaciones de R3.

## 6.2. Evaluación multimodal con U-Net clásica

Con el fin de analizar la efectividad de la fusión temprana de las modalidades, primero se utilizó una arquitectura U-Net con un único canal de procesamiento, ajustándose con respecto a las observaciones de ADC o DWI. Estos modelos, consistentes con los utilizados para los experimentos unimodales, fueron adaptados para recibir entradas multimodales mediante concatenaciones en la dimensión de los canales. Las métricas de Dice score y HD, basadas en las anotaciones de expertos para cada modalidad, están detalladas en la Tabla 3.

La Tabla 3 evidencia variaciones significativas en los Dice score y HD de los modelos multimodales, comparados con sus contrapartes unimodales, para diferentes combinaciones de entradas y anotaciones. Específicamente, los modelos que incluyen ADC en su configuración

**Tabla 3***Métricas computacionales para modelos U-Net clásicos.*

Mascara de Evaluacion	Entradas			R1		R2		R3	
	ADC	DWI	FLAIR	Dice $\uparrow$	HD $\downarrow$	Dice $\uparrow$	HD $\downarrow$	Dice $\uparrow$	HD $\downarrow$
ADC	$\checkmark$	$\times$	$\checkmark$	0.39 $\pm$ 0.31	37.83 $\pm$ 42.48	0.43 $\pm$ 0.31	39.12 $\pm$ 42.7	0.39 $\pm$ 0.28	46.38 $\pm$ 42.68
	$\checkmark$	$\checkmark$	$\times$	<b>0.53<math>\pm</math>0.29</b>	<b>27.2<math>\pm</math>36.99</b>	<b>0.54<math>\pm</math>0.28</b>	<u>37.92<math>\pm</math>44.12</u>	<b>0.49<math>\pm</math>0.26</b>	<b>38.83<math>\pm</math>40.27</b>
	$\times$	$\checkmark$	$\checkmark$	0.47 $\pm$ 0.32	<u>33.89<math>\pm</math>38.31</u>	0.48 $\pm$ 0.31	<b>37.33<math>\pm</math>42.99</b>	0.44 $\pm$ 0.29	<u>43.36<math>\pm</math>40.46</u>
	$\checkmark$	$\checkmark$	$\checkmark$	<u>0.5<math>\pm</math>0.3</u>	37.29 $\pm$ 39.83	<u>0.49<math>\pm</math>0.3</u>	41.98 $\pm$ 39.23	<u>0.47<math>\pm</math>0.27</u>	46.48 $\pm$ 41.02
DWI	$\checkmark$	$\times$	$\checkmark$	0.41 $\pm$ 0.3	44.19 $\pm$ 35.68	0.4 $\pm$ 0.31	55.03 $\pm$ 39.82	0.41 $\pm$ 0.29	62.79 $\pm$ 43.19
	$\checkmark$	$\checkmark$	$\times$	<b>0.61<math>\pm</math>0.28</b>	<u>30.48<math>\pm</math>27.73</u>	<b>0.57<math>\pm</math>0.29</b>	45.31 $\pm$ 37.44	<b>0.57<math>\pm</math>0.27</b>	57.24 $\pm$ 45.99
	$\times$	$\checkmark$	$\checkmark$	<u>0.59<math>\pm</math>0.28</u>	<b>25.58<math>\pm</math>25.97</b>	<u>0.55<math>\pm</math>0.28</u>	<b>42.65<math>\pm</math>37.59</b>	0.56 $\pm$ 0.27	<b>51.12<math>\pm</math>46.69</b>
	$\checkmark$	$\checkmark$	$\checkmark$	0.58 $\pm$ 0.27	32.11 $\pm$ 29.48	0.55 $\pm$ 0.29	<u>44.35<math>\pm</math>34.56</u>	<u>0.56<math>\pm</math>0.26</u>	<u>55.39<math>\pm</math>45.51</u>

*Nota.* Los mejores resultados según las anotaciones de ADC y DWI de cada experto están destacados en negrita, y los segundos mejores, subrayados.

de entrada muestran una mejora en el Dice score de 0.10 y 0.16 y una disminución del HD de 10.70 y 14.21 para las anotaciones de ADC y DWI, respectivamente. De la misma manera, los modelos que incluyen FLAIR en la entrada muestran mejoras de 0.41 y 0.39 en el Dice score y de 51.80 y 36.81 en HD para las mismas anotaciones. En contraste, los modelos que incorporan DWI en la entrada exhiben un cambio menor, con un decrecimiento de 0.02 para ADC y un incremento de 0.02 para DWI en el Dice score e incrementos de 2.51 y 7.22 para HD, con respecto a las anotaciones de ADC y DWI. Por otro lado, se continúan observando resultados superiores de DWI con respecto a ADC en Dice score por un promedio de 0.075 y en HD la superioridad es de los valores de ADC por un 5.33. La combinación de ADC y DWI obtuvo los mejores resultados según las anotaciones sobre ADC, excepto en HD para el R2, donde la combinación DWI y FLAIR es superior. Además, esta combinación ADC-DWI también lideró en términos de Dice score para las anotaciones en DWI, aunque en términos de HD, fue superada por la configuración que incluye DWI y FLAIR para todas las anotaciones de expertos. Sorprendentemente, el modelo que integra todas las modalidades disponibles no alcanzó el mejor desempeño en ninguna métrica, situándose segundo en Dice score para las anotaciones sobre ADC y las de DWI de R3, así como en HD para las anotaciones de DWI de R2. También es importante destacar que, aunque la combinación de FLAIR y ADC no supera las configuraciones que incluyen DWI, sí mejora el desempeño comparado con los modelos unimodales que usan ADC o FLAIR por separado.

**Tabla 4***Métricas médicas para modelos U-Net clásicos*

Mascara de Evaluacion	Entradas			R1			R2			R3		
	ADC	DWI	FLAIR	AVD ↓	ALD ↓	LF1 ↑	AVD ↓	ALD ↓	LF1 ↑	AVD ↓	ALD ↓	LF1 ↑
ADC	✓	×	✓	16.95±81.06	1.11±1.41	0.55±0.37	16.6±81.2	1.28±1.98	0.58±0.36	17.49±81.13	2.78±3.39	0.51±0.32
	✓	✓	×	<b>5.25±7.49</b>	<b>0.83±1.05</b>	<b>0.7±0.34</b>	<b>4.75±6.01</b>	<b>1.11±1.68</b>	<b>0.68±0.31</b>	<u>6.46±8.13</u>	<b>2.35±3.08</b>	<b>0.65±0.3</b>
	×	✓	✓	<u>5.53±9.85</u>	<u>1.09±1.31</u>	0.59±0.39	<u>5.82±9.39</u>	<u>1.19±1.81</u>	0.6±0.38	<b>6.4±9.45</b>	<u>2.39±2.96</u>	0.57±0.36
	✓	✓	✓	5.61±9.53	1.17±1.32	<u>0.63±0.34</u>	6.26±10.49	1.33±1.73	<u>0.61±0.34</u>	6.56±10.24	2.57±3.08	<u>0.58±0.32</u>
DWI	✓	×	✓	8.14±14.36	2.2±2.86	0.48±0.36	8.92±14.63	<u>2.09±2.76</u>	0.45±0.36	8.47±13.44	3.31±3.45	0.41±0.3
	✓	✓	×	<u>6.18±7.79</u>	2.5±3.56	0.64±0.33	<u>5.74±7.25</u>	2.65±3.89	0.54±0.34	<u>6.0±6.54</u>	3.35±3.98	0.53±0.31
	×	✓	✓	<b>5.05±8.01</b>	<b>1.63±2.64</b>	<b>0.67±0.35</b>	<b>5.09±7.88</b>	<b>1.52±2.1</b>	<b>0.6±0.36</b>	<b>4.87±8.09</b>	<b>2.52±3.15</b>	<u>0.58±0.33</u>
	✓	✓	✓	6.29±7.81	<u>2.07±2.88</u>	<u>0.66±0.33</u>	5.92±7.56	2.26±2.99	<u>0.56±0.35</u>	6.2±7.24	<u>2.81±3.16</u>	<b>0.58±0.31</b>

*Nota.* Los mejores resultados según las anotaciones de ADC y DWI de cada experto están destacados en negrita, y los segundos mejores, subrayados.

La Tabla 4 presenta los resultados de las métricas clínicas para diversos modelos que incluyen entradas multimodales a partir de una concatenación. La combinación de ADC y DWI mostró los mejores resultados, manteniendo consistencia en las evaluaciones de los tres radiólogos. En contraste, el modelo que usó ADC y FLAIR fue menos efectivo, particularmente en la estimación del volumen de la lesión, aunque esta diferencia fue menor al compararla con las máscaras de DWI. Por otro lado, la combinación de DWI y FLAIR fue superior en la estimación del volumen de la lesión y en la detección de lesiones independientes, incluso superando a los modelos que usaron solo DWI, destacando el beneficio de añadir FLAIR para la segmentación. La combinación de ADC y FLAIR tuvo un desempeño inferior en AVD comparada con otros modelos y con el uso exclusivo de ADC, pero la diferencia no fue tan notable contra las máscaras de ADC. Para las máscaras de ADC, la combinación de ADC y DWI logró generalizar mejor, mientras que para las de DWI, la combinación de DWI y FLAIR mostró la mejor efectividad.

### 6.3. Evaluación multimodal con base en la Dual U-Net y la Fusión temprana

En otro experimento, se evaluó la Dual U-Net, un modelo multimodal entrenado con distintas configuraciones de entrada en sus ramas convolucionales, ajustadas según las anotaciones de ADC y DWI. Esta arquitectura se adaptó para aceptar, en cualquiera de sus ramas convolucionales, concatenaciones de diversas combinaciones de imágenes del conjunto de datos, explorando así la contribución del FLAIR. Las métricas de Dice score y HD, obtenidas bajo la

comparación con anotaciones de expertos para cada modalidad, se presentan en la Tabla 5.

**Tabla 5**

*Métricas computacionales para modelos Dual U-Net usando imágenes de FLAIR como modalidad auxiliar mediante fusión temprana*

Mascara de Evaluacion	Entradas		R1		R2		R3	
	Codificador 1 (B1)	Codificador 2 (B2)	Dice $\uparrow$	HD $\downarrow$	Dice $\uparrow$	HD $\downarrow$	Dice $\uparrow$	HD $\downarrow$
ADC	ADC-DWI	FLAIR	<b>0.53<math>\pm</math>0.3</b>	<u>32.16<math>\pm</math>37.66</u>	<u>0.52<math>\pm</math>0.29</u>	<b>37.53<math>\pm</math>39.64</b>	0.47 $\pm$ 0.27	<u>43.72<math>\pm</math>43.29</u>
	ADC-FLAIR	DWI-FLAIR	0.49 $\pm$ 0.31	35.82 $\pm$ 43.34	0.52 $\pm$ 0.29	42.28 $\pm$ 47.78	0.46 $\pm$ 0.28	47.96 $\pm$ 44.4
	ADC	DWI	<u>0.52<math>\pm</math>0.27</u>	<b>28.17<math>\pm</math>36.38</b>	<b>0.52<math>\pm</math>0.25</b>	<u>39.01<math>\pm</math>43.89</u>	<b>0.48<math>\pm</math>0.25</b>	<b>39.74<math>\pm</math>39.05</b>
	ADC	FLAIR-DWI	0.52 $\pm$ 0.29	35.08 $\pm$ 38.62	0.52 $\pm$ 0.28	40.12 $\pm$ 37.84	<u>0.47<math>\pm</math>0.27</u>	44.44 $\pm$ 39.58
DWI	ADC-DWI	FLAIR	0.51 $\pm$ 0.31	34.04 $\pm$ 36.28	0.5 $\pm$ 0.31	43.57 $\pm$ 40.33	0.48 $\pm$ 0.3	60.17 $\pm$ 49.98
	ADC-FLAIR	DWI-FLAIR	<u>0.57<math>\pm</math>0.3</u>	<u>33.58<math>\pm</math>33.43</u>	0.55 $\pm$ 0.3	<u>39.12<math>\pm</math>36.34</u>	0.56 $\pm$ 0.29	<b>53.48<math>\pm</math>48.71</b>
	ADC	DWI	<b>0.62<math>\pm</math>0.28</b>	<b>28.47<math>\pm</math>28.08</b>	<b>0.56<math>\pm</math>0.29</b>	40.45 $\pm$ 34.63	<b>0.58<math>\pm</math>0.27</b>	56.9 $\pm$ 47.78
	ADC	FLAIR-DWI	0.57 $\pm$ 0.31	33.68 $\pm$ 32.73	<u>0.55<math>\pm</math>0.3</u>	<b>38.1<math>\pm</math>31.02</b>	<u>0.56<math>\pm</math>0.29</u>	<u>56.61<math>\pm</math>50.26</u>

*Nota.* Los mejores resultados según las anotaciones de ADC y DWI de cada experto están destacados en negrita, y los segundos mejores, subrayados.

La Tabla 5 muestra variabilidad en el rendimiento entre las distintas configuraciones analizadas, con excepción de la configuración que excluye FLAIR. No obstante, se continúa evidenciando un aumento significativo en las métricas comparadas con aquellas de las configuraciones unimodales. Adicionalmente, se nota un aumento promedio del 0.13 y 0.005 en el Dice score y del 6.23 y 3.41 en HD en comparación con configuraciones unimodales que utilizan U-Nets clásicas para ADC y DWI respectivamente. Particularmente, el modelo que incluye en las ramas codificadoras B1: ADC y B2: DWI consigue consistentemente los mejores resultados de solapamiento y definición de bordes. A su vez, la combinación B1: ADC-DWI y B2: FLAIR también alcanza resultados competitivos con las anotaciones de ADC, aunque este rendimiento no se mantiene al analizar las anotaciones en DWI. Inesperadamente, la configuración con FLAIR en ambas ramas codificadoras (B1: ADC-FLAIR, B2: DWI-FLAIR) muestra competitividad para las anotaciones de DWI, pero su rendimiento disminuye considerablemente en comparación con otras configuraciones para las anotaciones de ADC.

Como se evidencia en la Tabla 6, en contraste con las métricas Dice y HD de la Tabla 5, el uso de FLAIR mejora la predicción del volumen absoluto de las lesiones en imágenes de ADC, logrando el mejor desempeño en AVD en todos los experimentos. Particularmente, la configuración B1:ADC, B2:FLAIR-DWI consistentemente ofrece predicciones de volumen de la lesión más cercanas a las anotaciones de los tres radiólogos y mantiene un rendimiento

**Tabla 6**

*Métricas médicas para modelos Dual U-Net usando imágenes de FLAIR como modalidad auxiliar mediante fusión temprana*

Mascara de Evaluacion	Entradas		R1			R2			R3		
	Codificador 1 (B1)	Codificador 2 (B2)	AVD ↓	ALD ↓	LF1 ↑	AVD ↓	ALD ↓	LF1 ↑	AVD ↓	ALD ↓	LF1 ↑
ADC	ADC-DWI	FLAIR	5.94±9.04	<b>0.96±1.19</b>	<u>0.66±0.36</u>	<u>5.74±8.31</u>	<u>1.13±1.77</u>	0.63±0.34	6.75±9.25	<u>2.33±3.07</u>	<u>0.61±0.32</u>
	ADC-FLAIR	DWI-FLAIR	21.35±87.69	<u>0.98±1.27</u>	0.63±0.38	22.3±87.67	1.15±1.58	<u>0.64±0.35</u>	22.77±87.49	2.5±2.91	0.57±0.32
	ADC	DWI	<u>5.52±9.87</u>	1.0±1.2	<b>0.69±0.34</b>	5.79±8.68	<b>1.06±1.41</b>	<b>0.68±0.31</b>	<u>6.32±9.55</u>	2.44±2.68	<b>0.62±0.29</b>
	ADC	FLAIR-DWI	<b>4.84±7.39</b>	1.46±1.74	0.65±0.33	<b>4.85±7.9</b>	1.19±1.52	0.64±0.33	<b>5.71±7.1</b>	<b>2.31±2.5</b>	0.57±0.31
DWI	ADC-DWI	FLAIR	5.58±9.18	1.69±2.62	0.62±0.38	5.83±10.49	1.61±2.23	0.58±0.37	6.26±10.99	2.72±3.24	0.52±0.33
	ADC-FLAIR	DWI-FLAIR	7.33±23.5	1.5±2.37	0.65±0.35	7.89±23.38	<u>1.35±1.88</u>	0.59±0.36	7.26±23.39	<u>2.31±3.06</u>	0.55±0.33
	ADC	DWI	<b>3.62±5.48</b>	<u>1.48±2.14</u>	<b>0.69±0.3</b>	<b>4.59±6.25</b>	1.56±1.86	<u>0.59±0.33</u>	<b>4.68±6.3</b>	2.41±2.75	<u>0.57±0.3</u>
	ADC	FLAIR-DWI	<u>4.46±8.98</u>	<b>1.33±2.1</b>	<u>0.66±0.33</u>	<u>5.4±10.43</u>	<b>1.15±1.24</b>	<b>0.61±0.34</b>	<u>5.59±11.06</u>	<b>2.0±2.51</b>	<b>0.57±0.32</b>

*Nota.* Los mejores resultados según las anotaciones de ADC y DWI de cada experto están destacados en negrita, y los segundos mejores, subrayados.

competitivo con las máscaras de DWI. Además, la configuración B1:ADC y B2:DWI destaca notablemente en la predicción de la cantidad de lesiones, abarcando ambas máscaras y todas las evaluaciones de los radiólogos.

#### 6.4. Evaluación multimodal con base en la Dual U-Net y la Fusión tardía

Como parte de este trabajo también se evaluó el modelo que combina secuencias radiológicas mediante ramas específicas, que realiza una fusión tardía de características en el espacio embebido. Similar a la Dual U-Net con fusión temprana, este modelo se calibró de acuerdo con los hallazgos radiológicos delineados en las secuencias de ADC y DWI. Las métricas de Dice score y HD, derivadas de comparar con anotaciones de expertos para cada modalidad, están detalladas en la Tabla 7.

**Tabla 7**

*Métricas computacionales para modelos Dual U-Net usando imágenes de FLAIR como modalidad auxiliar mediante fusión tardía.*

Mascara de Evaluacion	Entradas	R1		R2		R3	
		Dice ↑	HD ↓	Dice ↑	HD ↓	Dice ↑	HD ↓
ADC	ADC;DWI;FLAIR	0.51±0.29	33.52±37.6	0.51±0.28	38.78±38.98	0.46±0.26	42.95±40.18
DWI	ADC;DWI;FLAIR	0.61±0.27	35.68±30.48	0.54±0.29	47.7±38.46	0.57±0.27	63.07±47.98

*Nota.* Los mejores resultados según las anotaciones de ADC y DWI de cada experto están destacados en negrita, y los segundos mejores, subrayados.

Como se muestra en la tabla 7, el método, que realiza una fusión tardía de las caracterís-

ticas profundas, logra superar las configuraciones unimodales en las métricas Dice score y HD. Los valores de Dice score y HD indican que la estimación de la lesión apoyada por imágenes de FLAIR mediante una rama independiente no revela información adicional que el decodificador aproveche para estimar lesiones con un mayor solapamiento o con una mejor definición de bordes, con respecto a la fusión temprana. Más específicamente, se evidenció una desmejora promedio para todos los expertos de 0.07 y 0.1 en Dice score en ADC y DWI, y 1.11 en HD para ADC, comparando con las anotaciones en ADC y DWI respectivamente.

**Tabla 8**

*Métricas médicas para modelos Dual U-Net usando imágenes de FLAIR como modalidad auxiliar mediante fusión tardía*

Mascara de Evaluacion	Entradas	R1			R2			R3		
		AVD ↓	ALD ↓	LF1 ↑	AVD ↓	ALD ↓	LF1 ↑	AVD ↓	ALD ↓	LF1 ↑
ADC	ADC;DWI;FLAIR	<b>6.42±9.54</b>	<b>1.04±1.09</b>	<b>0.64±0.34</b>	<u>5.95±8.74</u>	<u>1.13±1.43</u>	<u>0.63±0.32</u>	<u>6.7±9.48</u>	<u>2.44±2.64</u>	<u>0.59±0.29</u>
DWI	ADC;DWI;FLAIR	<b>4.96±6.2</b>	<b>3.04±4.35</b>	<b>0.62±0.32</b>	<u>4.86±5.42</u>	<u>3.07±4.5</u>	<u>0.52±0.34</u>	<u>4.78±5.81</u>	<u>3.59±4.24</u>	<u>0.52±0.3</u>

*Nota.* Los mejores resultados según las anotaciones de ADC y DWI de cada experto están destacados en negrita, y los segundos mejores, subrayados.

Como sugieren los valores mostrados en la tabla 8 y al compararse con la fusión temprana en la tabla 6 la predicción realizada por la fusión tardía al momento de predecir el volumen de la lesión y la cantidad de lesiones existentes no logra igualar en ninguna métrica valorada al método alternativo. Aun así, como método independiente, la generalización entre radiólogos es robusta al observar poca variación ante las diferentes marcaciones.

## 7. CONCLUSIONES Y TRABAJO FUTURO

En este trabajo se implementó una arquitectura multimodal para soportar la segmentación de lesiones isquémicas, utilizando observaciones ADC y DWI, así como también valorando la inclusión del mapa paramétrico FLAIR. El método incluye tareas de múltiple segmentación, incluyendo el ajuste con respecto a máscaras anotadas de forma independiente en ADC y DWI. Bajo esta premisa, y considerando una arquitectura base, se exploraron diferentes alternativas de fusión. La primera configuración utilizó una arquitectura Dual UNet que cuenta con dos codificadores y decodificadores. Esta arquitectura fue adaptada para recibir imágenes de ADC y DWI en la entrada, soportadas por imágenes FLAIR concatenadas para llevar a cabo una fusión temprana en cualquiera de las ramas de procesamiento. Por otro lado, la segunda arquitectura utilizó una rama dedicada para cada modalidad del conjunto de datos, dando como resultado una red con tres codificadores que terminaban en una fusión tardía de las características más relevantes del ACV isquémico. Desde este único embebido, se desplegaban dos decodificadores para estimar los hallazgos clínicos de las modalidades ADC y DWI. Cada arquitectura desarrollada fue utilizada para estimar las lesiones de ACV isquémico sobre un conjunto de pacientes de prueba no vistos durante el entrenamiento.

En los resultados obtenidos, se evidenció que las configuraciones que realizan una fusión de características multimodales de manera temprana y tardía, superaron significativamente las configuraciones unimodales, mostrando el potencial de estas arquitecturas para extraer información multicontexto y explotar la complementariedad de las secuencias utilizadas en el protocolo imagenológico de diagnóstico del ACV. Particularmente, se resalta la mejora observada en las métricas clínicas, mostrando una disminución en la diferencia de segmentos de lesiones estimadas por estudio y un incremento en la detección de estos. Esto es relevante, ya que muestra que un análisis multimodal tiene potencial para estimar lesiones con patrones variados, factor

importante para determinar la etiología del ACV <sup>2627</sup>.

Adicionalmente, en la comparación con la línea base multimodal, se observaron rendimientos comparables de las mejores configuraciones para cada arquitectura. No obstante, se puede observar que ajustar las arquitecturas desarrolladas da en promedio modelos con mayor capacidad para segmentar las lesiones de ACV (Dice promedio UNet Clásica=0.51, Fusión temprana=0.53, Fusión tardía=0.56 y AVD promedio UNet Clásica=7.38mL, Fusión temprana=7.64mL, Fusión tardía=5.69mL). En general, las mejores configuraciones fueron obtenidas a partir de utilizar exclusivamente el ADC y DWI, lo cual concuerda con los resultados del estado del arte <sup>27</sup>. Con respecto al aporte del FLAIR, las configuraciones que lo incluyeron junto al ADC y DWI, obtuvieron resultados comparables pero usualmente más bajos. Esto puede deberse a que el FLAIR muestra hallazgos radiológicos después de la ventana de tratamiento trombolítico (después de 6 horas) <sup>28</sup>, dando como resultado la ausencia de patrones visuales de ACV para los casos dentro de esta ventana temporal (menor a 6 horas).

Como perspectivas de este trabajo, será necesario realizar una evaluación estratificada basada en el tiempo transcurrido desde el inicio de los síntomas para analizar detalladamente el aporte del FLAIR. Alternativamente, se podría evaluar la inclusión del tiempo en el análisis multimodal con aprendizaje profundo, determinando la relevancia de los hallazgos de cada secuencia radiológica dentro de ventanas temporales definidas. Además, sería provechoso explorar alternativas computacionales a los módulos de atención cruzada aditiva, mejorando la fusión de características en diversos niveles de procesamiento.

---

<sup>26</sup> Hans-Christian KOENNECKE et al. “Scattered brain infarct pattern on diffusion-weighted magnetic resonance imaging in patients with acute ischemic stroke”. In: *Cerebrovascular Diseases* 11.3 (2001), pp. 157–163.

<sup>27</sup> Ezequiel de la ROSA et al. “A Robust Ensemble Algorithm for Ischemic Stroke Lesion Segmentation: Generalizability and Clinical Utility Beyond the ISLES Challenge”. In: *arXiv preprint arXiv:2403.19425* (2024).

<sup>28</sup> A MESHKSAR et al. “Role of EPI-FLAIR in patients with acute stroke: a comparative analysis with FLAIR”. in: *American Journal of Neuroradiology* 35.5 (2014), pp. 878–883.

## BIBLIOGRAFÍA

ASHTARI, Pooya et al. “Factorizer: A scalable interpretable approach to context modeling for medical image segmentation”. In: *Medical image analysis* 84 (2023), p. 102706 (cit. on p. 21).

BAL, Abhishek et al. “A robust ischemic stroke lesion segmentation technique using two-pathway 3D deep neural network in MR images”. In: *Multimedia Tools and Applications* (Oct. 2023). DOI: 10.1007/s11042-023-16689-9 (cit. on p. 12).

*Caplan’s Stroke*. Aug. 5, 2016. DOI: 10.1017/cbo9781316095805 (cit. on p. 13).

CLÈRIGUES, Albert et al. “Acute and sub-acute stroke lesion segmentation from multimodal MRI”. In: *Computer Methods and Programs in Biomedicine* 194 (2020). DOI: 10.1016/j.cmpb.2020.105521. arXiv: 1810.13304 (cit. on p. 20).

FEIGIN, Valery L et al. “World Stroke Organization (WSO): global stroke fact sheet 2022”. In: *International Journal of Stroke* 17.1 (2022), pp. 18–29 (cit. on pp. 10, 13).

FESKE, Steven K. “Ischemic Stroke”. In: *The American Journal of Medicine* 134.12 (Dec. 2021), pp. 1457–1464. DOI: 10.1016/j.amjmed.2021.07.027 (cit. on p. 13).

GÓMEZ, Santiago et al. “Ischemic Stroke Segmentation from a Cross-Domain Representation in Multimodal Diffusion Studies”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2023, pp. 776–785 (cit. on p. 12).

KAMNITSAS, Konstantinos et al. “Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation”. In: *Medical Image Analysis* 36 (Feb. 2017), pp. 61–78. DOI: 10.1016/j.media.2016.10.004 (cit. on p. 20).

KOENNECKE, Hans-Christian et al. “Scattered brain infarct pattern on diffusion-weighted magnetic resonance imaging in patients with acute ischemic stroke”. In: *Cerebrovascular Diseases* 11.3 (2001), pp. 157–163 (cit. on p. 47).

LEE, Seung-Hoon. *Stroke Revisited: Diagnosis and treatment of ischemic stroke*. Jan. 2017. DOI: 10.1007/978-981-10-1424-6 (cit. on p. 10).

LI, Xiangrui et al. “The first step for neuroimaging data analysis: DICOM to NIfTI conversion”. In: *Journal of neuroscience methods* 264 (2016), pp. 47–56 (cit. on p. 32).

LIU, Liangliang et al. “Attention convolutional neural network for accurate segmentation and quantification of lesions in ischemic stroke disease”. In: *Medical Image Analysis* 65 (2020), p. 101791 (cit. on p. 21).

MARTEL, Anne L et al. “Measurement of infarct volume in stroke patients using adaptive segmentation of diffusion weighted MR images”. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI’99: Second International Conference, Cambridge, UK, September 19-22, 1999. Proceedings 2*. Springer. 1999, pp. 22–31 (cit. on p. 11).

MESHKSAR, A et al. “Role of EPI-FLAIR in patients with acute stroke: a comparative analysis with FLAIR”. In: *American Journal of Neuroradiology* 35.5 (2014), pp. 878–883 (cit. on p. 47).

NAG, Manas Kumar et al. “Computer-assisted delineation of cerebral infarct from diffusion-weighted MRI using Gaussian mixture model”. In: *International Journal of Computer Assisted Radiology and Surgery* 12.4 (Jan. 2017), pp. 539–552. DOI: 10.1007/s11548-017-1520-x (cit. on p. 11).

NEUMANN, Anders B et al. “Interrater agreement for final infarct MRI lesion delineation”. In: *Stroke* 40.12 (2009), pp. 3768–3771 (cit. on p. 11).

OU, Yanglan et al. *LambdaUNET: 2.5D stroke lesion segmentation of Diffusion-Weighted MR Images*. Jan. 2021, pp. 731–741. DOI: 10.1007/978-3-030-87193-2\_69 (cit. on p. 21).

PINTO, Adriano et al. “Combining unsupervised and supervised learning for predicting the final stroke lesion”. In: *Medical image analysis* 69 (2021), p. 101888 (cit. on p. 22).

QI, Kehan et al. *X-NET: Brain stroke lesion segmentation based on depthwise Separable convolution and Long-Range Dependencies*. Jan. 2019, pp. 247–255. DOI: 10.1007/978-3-030-32248-9\_28 (cit. on p. 19).

ROSA, Ezequiel de la et al. “A Robust Ensemble Algorithm for Ischemic Stroke Lesion Segmentation: Generalizability and Clinical Utility Beyond the ISLES Challenge”. In: *arXiv preprint arXiv:2403.19425* (2024) (cit. on p. 47).

SHIN, Hyunkwang et al. “Automated segmentation of chronic stroke lesion using efficient U-Net architecture”. In: *Biocybernetics and Biomedical Engineering* 42.1 (2022), pp. 285–294. DOI: <https://doi.org/10.1016/j.bbe.2022.01.002> (cit. on pp. 11, 19).

SILVERA, Stéphane et al. “Spontaneous Intracerebral Hematoma on Diffusion-weighted Images: Influence of T2-shine-through and T2-blackout Effects”. In: *American Journal of Neuroradiology* 26.2 (2005), pp. 236–241. eprint: <https://www.ajnr.org/content/26/2/236.full.pdf> (cit. on p. 11).

STRANDGAARD, S and PAULSON, OB. “Pathophysiology of stroke”. In: *Journal of cardiovascular pharmacology* 15 (1990), S38–S42 (cit. on p. 13).

WANG, Shuqiang et al. “Brain stroke lesion segmentation using consistent perception generative adversarial network”. In: *Neural Computing and Applications* 34.11 (Mar. 2022), pp. 8657–8669. DOI: 10.1007/s00521-021-06816-8 (cit. on pp. 11, 19).

WOODRUFF, Trent M. et al. “Pathophysiology, treatment, and animal and cellular models of human ischemic stroke”. In: *Molecular Neurodegeneration* 6.1 (Jan. 2011). DOI: 10.1186/1750-1326-6-11 (cit. on p. 13).

ZHANG, Rongzhao et al. “Automatic Segmentation of Acute Ischemic Stroke From DWI Using 3-D Fully Convolutional DenseNets”. In: *IEEE Transactions on Medical Imaging* 37.9 (2018), pp. 2149–2160. DOI: 10.1109/TMI.2018.2821244 (cit. on p. 12).

ZHANG, Yue et al. *A 3D+2D CNN approach incorporating boundary loss for stroke lesion segmentation*. Jan. 2020, pp. 101–110. DOI: 10.1007/978-3-030-59861-7\_11 (cit. on p. 20).

ZHOU, Yongjin et al. “D-UNET: a Dimension-Fusion U shape network for chronic stroke lesion segmentation”. In: *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 18.3 (May 2021), pp. 940–950. DOI: 10.1109/tcbb.2019.2939522 (cit. on p. 19).