



MODELO GENERATIVO PARA LA SÍNTESIS DE MODALIDADES RADIOLÓGICAS
CON ÉNFASIS EN LA PRESERVACIÓN DE PATRONES VISUALES DE LESIONES DE
ACCIDENTE CEREBROVASCULAR ISQUÉMICO

GYSELIS VÁSQUEZ GARCÉS

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISICOMECAÑICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA

2024

MODELO GENERATIVO PARA LA SÍNTESIS DE MODALIDADES RADIOLÓGICAS
CON ÉNFASIS EN LA PRESERVACIÓN DE PATRONES VISUALES DE LESIONES DE
ACCIDENTE CEREBROVASCULAR ISQUÉMICO

GYSSELIS VÁSQUEZ GARCÉS

Trabajo de Grado presentado en cumplimiento de los requisitos para optar al título de:
Ingeniera de Sistemas

Director:

Fabio Martínez Carrillo

Doctor en Ingeniería de Sistemas y Computación

Codirector:

Gustavo Adolfo Garzón Villamizar

Magíster en Ingeniería de Sistemas e Informática

Codirector:

Santiago Gómez Hernández

Magíster en Ingeniería de Sistemas e Informática

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISICOMECAÑICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA

2024

DEDICATORIA

A mis padres, Yaniris y Cristóbal, y a mi hermano Jhonatan, quienes han sido mi constante apoyo y motivación en cada paso del camino.

También, con todo mi amor, a mi querida gatita Micho, y en memoria de mis angelitos, Pepe y Kitty, que han llenado de alegría y consuelo mis días.

AGRADECIMIENTOS

Quiero expresar mi más sincero agradecimiento a mi director, el profesor Fabio, y a mis codirectores, Santiago y Gustavo, por su invaluable guía y apoyo durante este proceso. Su dedicación y consejos han sido fundamentales para alcanzar este logro. Les tengo una profunda admiración por su compromiso y paciencia.

También quiero destacar el cariño y el apoyo constante de mis compañeros del grupo de investigación *BIVL²ab*, quienes han compartido su conocimiento y han estado siempre dispuestos a ayudarme.

Agradezco infinitamente a mis padres y a mi hermano por creer siempre en mí y apoyarme en cada etapa de mi vida. Su amor incondicional ha sido mi fortaleza.

No puedo dejar de agradecer a mi mejor amiga, Laura Carolina, por estar siempre a mi lado, inspirándome a superarme y perseguir mis sueños, a pesar de las barreras de la distancia.

Finalmente, estoy agradecida con mis amigos por acompañarme en este viaje, por ser el hombro en el que he podido llorar y por ser esa voz de aliento en los momentos en que más lo necesitaba.

CONTENIDO

	pág.
INTRODUCCIÓN	11
1. FUNDAMENTOS Y TRABAJOS PREVIOS	15
1.1. ACV isquémico y las imágenes médicas para soportar el diagnóstico	15
1.2. Estrategias de aprendizaje profundo para el aprendizaje generativo	19
1.3. Mecanismos generativos y síntesis de imágenes en ACV	25
2. PROBLEMA DE INVESTIGACIÓN	29
3. OBJETIVOS	31
3.1. Objetivo general	31
3.2. Objetivos específicos	31
4. MÉTODO PROPUESTO	32
4.1. Generador: una arquitectura codificador-decodificador	33
4.2. Discriminador PatchGAN	34
4.3. Optimización de la generación por tipo de tejido	35
4.3.1 Pesos de clase binarios.	35
4.3.2 Pesos de clase dilatados.	35
5. DISEÑO EXPERIMENTAL	37
5.1. Datos	37
5.2. Preprocesamiento	37
5.3. Configuración de la arquitectura	38
5.4. Validación	39

5.4.1 Métricas de evaluación.	39
5.4.2 Prueba estadística.	42
6. EVALUACIÓN Y RESULTADOS	43
6.1. Secuencias FLAIR sintéticas como soporte en el análisis de DWI-FLAIR mismatch	43
6.2. DWI y mapas ADC sintéticos	48
7. CONCLUSIONES Y TRABAJO FUTURO	53
BIBLIOGRAFÍA	56

LISTA DE FIGURAS

	pág.
Figura 1. NCCT de dos pacientes diferentes	16
Figura 2. DWI y NCCT de un mismo paciente	18
Figura 3. DWI y ADC de un mismo paciente	18
Figura 4. DWI-FLAIR Mismatch	20
Figura 5. Modelo generativo adversario (GAN).	21
Figura 6. Síntesis de RM a partir de TC (TC → RM)	22
Figura 7. Modelo Pix2Pix.	24
Figura 8. Arquitectura de modelo CycleGAN.	25
Figura 9. Esquema del enfoque propuesto.	32
Figura 10. Aplicación de métodos de ponderación	36
Figura 11. Resultados cualitativos de las secuencias FLAIR sintéticas a partir de NCCT.	43
Figura 12. Resultados cualitativos de las secuencias FLAIR sintéticas a partir de DWI.	44
Figura 13. Resultados cualitativos de las secuencias FLAIR sintéticas a partir de mapas ADC.	45
Figura 14. Resultados cualitativos de mapas ADC sintéticos a partir de NCCT.	48
Figura 15. Resultados cualitativos de secuencias DWI sintéticas a partir de NCCT.	49

LISTA DE TABLAS

	pág.
Tabla 1. Distribución de pacientes por pares de modalidades de imagen.	39
Tabla 2. Métricas de evaluación de calidad estructural y perceptual de imágenes FLAIR sintéticas a partir de NCCT, DWI y mapas ADC. La columna de pesos de clase (CW) corresponde a ninguno (N), binario (B) y dilatado (D).	46
Tabla 3. Métricas de evaluación de calidad estructural y perceptual de imágenes FLAIR sintéticas, delimitadas por la región de la lesión isquémica, a partir de NCCT, DWI y mapas ADC. La columna de pesos de clase (CW) corresponde a ninguno (N), binario (B) y dilatado (D).	47
Tabla 4. Evaluación de la calidad estructural y perceptual de mapas ADC/imágenes DWI sintéticas a partir de NCCT. La columna de pesos de clase (CW) corresponde a ninguno (N), binario (B) y dilatado (D).	50
Tabla 5. Evaluación de la calidad estructural y perceptual de mapas ADC/imágenes DWI sintéticas, delimitadas por la región de la lesión isquémica, a partir de NCCT. La columna de pesos de clase (CW) corresponde a ninguno (N), binario (B) y dilatado (D).	51

RESUMEN

TÍTULO: Modelo generativo para la síntesis de modalidades radiológicas con énfasis en la preservación de patrones visuales de lesiones de accidente cerebrovascular isquémico *

AUTOR: Gysselis Vásquez Garcés **

PALABRAS CLAVE: Accidente cerebrovascular isquémico, imágenes médicas sintéticas, preservación de lesiones, arquitecturas generativas, mapas de pesos de clase.

DESCRIPCIÓN: La evaluación rápida del accidente cerebrovascular isquémico es crucial para proporcionar un tratamiento oportuno y minimizar las secuelas neurológicas. El protocolo estándar comienza con una tomografía computarizada sin contraste (NCCT), que aunque rápida, ofrece un contraste deficiente entre el tejido lesionado y sano. Las secuencias de imagen ponderada por difusión (DWI) y FLAIR son más precisas para la caracterización de las lesiones. Sin embargo, su adquisición puede ser prolongada y dependiente de la disponibilidad del escáner de resonancia magnética (RM), que no siempre es inmediata. Este estudio propone un enfoque generativo para sintetizar no solo secuencias FLAIR a partir de NCCT, sino también DWI y mapas ADC, enfocándose en las regiones afectadas por el accidente cerebrovascular. Utilizando pesos de clase binarios y dilatados, este método mejora la representación de texturas y características estructurales de las lesiones isquémicas, manteniendo una alta fidelidad visual en todos los dominios de traducción. En un estudio retrospectivo de 98 pacientes con ACV isquémico, se lograron resultados prometedores: las secuencias FLAIR sintetizadas desde NCCT mostraron un SSIM de 0.90 y un SSIM delimitado por lesiones de 0.56. Similarmente, las traducciones a DWI y ADC alcanzaron un SSIM de 0.87 a nivel global y a nivel de lesión se reportaron SSIM de 0.58 en ADC y 0.55 en DWI, demostrando una alta correlación estructural con las secuencias originales. Este método puede reducir significativamente los retrasos en la adquisición de imágenes críticas, proporcionando una herramienta valiosa en entornos clínicos con acceso limitado a RM, mejorando así la rapidez y precisión en el diagnóstico y tratamiento de los ACV.

* Trabajo de investigación

** Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería de Sistemas e Informática. Director: Fabio Martínez, PhD en ingeniería de sistemas y computación, análisis de imágenes y análisis de vídeo.

ABSTRACT

TITLE: Generative model for synthesis of radiological modalities with emphasis on preservation of visual patterns of ischemic stroke lesions *

AUTHOR: Gysselis Vásquez Garcés **

KEYWORDS: Ischemic stroke, synthetic medical images, lesion preservation, generative architectures, class weights maps.

DESCRIPTION: The rapid evaluation of ischemic stroke is crucial to provide timely treatment and minimize neurological sequelae. The standard protocol begins with a non-contrast computed tomography (NCCT), which, although quick, offers poor contrast between injured and healthy tissue. Diffusion-weighted imaging (DWI) and FLAIR sequences are more accurate for lesion characterization. However, their acquisition can be prolonged and dependent on the availability of the magnetic resonance (MR) scanner, which is not always immediate. This study proposes a generative approach to synthesize not only FLAIR sequences from NCCT but also DWI and ADC maps, focusing on the regions affected by stroke. Using binary and dilated class weights, this method improves the representation of textures and structural features of ischemic lesions, maintaining high visual fidelity across all translation domains. In a retrospective study of 98 patients with ischemic stroke, promising results were achieved: the FLAIR sequences synthesized from NCCT showed an SSIM of 0.90 and a lesion-bounded SSIM of 0.56. Similarly, the translations to DWI and ADC reached a global SSIM of 0.87, and at the lesion level, SSIMs of 0.58 in ADC and 0.55 in DWI were reported, demonstrating high structural correlation with the original sequences. This method can significantly reduce delays in acquiring critical images, providing a valuable tool in clinical settings with limited access to MR, thereby improving the speed and accuracy of stroke diagnosis and treatment.

* Research work

** Faculty of Physics-Mechanics Engineering. School of Systems Engineering and Informatics. Advisor: Fabio Martínez Carrillo, PhD. Computer and systems engineering, medical image analysis and video analysis.

INTRODUCCIÓN

El accidente cerebrovascular (ACV) es una de las principales causas de muerte y discapacidad en el mundo ¹. El ACV es causado por la interrupción repentina del flujo sanguíneo al cerebro, ocasionando la muerte de células cerebrales con el transcurrir del tiempo. En particular, el ACV isquémico es causado por la oclusión de un vaso sanguíneo, siendo el tipo más común de ACV con aproximadamente el 87% de los casos ². Para minimizar sus consecuencias es vital un diagnóstico temprano y preciso que oriente el tratamiento más adecuado para cada paciente ¹.

El diagnóstico del ACV se basa principalmente en la caracterización de las lesiones cerebrales mediante estudios de imagenología como la tomografía computarizada (TC) y la resonancia magnética (RM). La TC sin contraste (NCCT, por sus siglas en inglés) suele ser el primer estudio que se realiza a pacientes con sospecha de ACV, ya que permite distinguir entre un ACV de tipo hemorrágico e isquémico. Además, esta secuencia permite una exploración rápida y cuenta con una alta disponibilidad en centros clínicos. Sin embargo, este estudio tiene una baja sensibilidad para caracterizar lesiones isquémicas con pocas horas desde el inicio de los síntomas (fase aguda), debido a que presenta un contraste deficiente entre el tejido normal e hipoperfundido ¹. Por su parte, la RM ofrece imágenes más detalladas y sensibles, con secuencias o ponderaciones que resaltan diferentes características de las lesiones. Las imágenes de RM ponderadas por difusión (DWI, por sus siglas en inglés) son cruciales para el diagnóstico de ACV isquémico agudo, ya

¹ Rodrigo PARDO TURRIAGO. “Guía de práctica clínica para el diagnóstico, tratamiento y rehabilitación del episodio agudo de ataque cerebrovascular en población mayor de 18 años: de las recomendaciones a su implementación”. In: *Acta Neurológica Colombiana* 31.4 (2015), pp. 462–467.

² Wei XIANG et al. “FLAIR vascular hyperintensity combined with asymmetrical prominent veins in acute anterior circulation ischemic stroke: prediction of collateral circulation and clinical outcome”. In: *European Journal of Medical Research* 28.1 (2023), p. 446.

que permiten detectar áreas con restricción del movimiento de las moléculas de agua, observadas como regiones hiperintensas en las imágenes ³. El coeficiente de difusión aparente (ADC, por sus siglas en inglés) complementa esta información proporcionando una medida cuantitativa del movimiento de las moléculas de agua dentro de los tejidos, mostrando las áreas isquémicas como regiones hipointensas ³. Complementariamente, se capturan secuencias FLAIR (Fluid-Attenuated Inversion Recovery), que permiten aumentar la visibilidad de la lesión y resaltar características distintivas del ACV debido a que suprime la señal del líquido cefalorraquídeo ^{4 2}. Esta secuencia es crucial en la estimación del tiempo transcurrido desde el inicio del ACV cuando este se desconoce. No obstante, la RM es un estudio costoso, que implica un tiempo de adquisición prolongado y tiene contraindicaciones para algunos pacientes ¹. Cada modalidad de imagenología tiene sus ventajas y ofrece complementariedad para el diagnóstico del ACV, aportando información sobre el estado del tejido cerebral. Además, para un pronóstico favorable del ACV un factor clave es el tiempo de atención, por lo cual la disposición de múltiples secuencias imagenológicas oportunamente es ponderado sobre otros factores.

Una alternativa para aproximar este problema desde un punto de vista computacional son los métodos de traducción entre distintas secuencias radiológicas con el fin de mejorar la calidad y la disponibilidad de las imágenes. Estos métodos usan modelos basados en aprendizaje profundo, especialmente redes generativas adversarias (GAN), que aprenden

³ Laura FILOGRANA et al. "Diffusion weighted imaging (DWI) and apparent diffusion coefficient (ADC) values for detection of malignant vertebral bone marrow lesions." In: *European Review for Medical & Pharmacological Sciences* 22.3 (2018).

⁴ Smitha MAKKAT et al. "Signs of acute stroke seen on fluid-attenuated inversion recovery MR imaging". In: *American Journal of Roentgenology* 179.1 (2002), pp. 237–243.

a traducir entre diferentes modalidades de imagenología, como TC y RM^{5 6}. Sin embargo, estos modelos presentan desafíos como la pérdida de información visual y la dificultad para integrar la información de la lesión en el proceso de traducción⁷.

En este trabajo, se presenta una estrategia generativa para la traducción entre secuencias radiológicas de estudios TC y RM de pacientes con ACV isquémico. Esta arquitectura incorpora una función de pérdida que permite evaluar la calidad de reconstrucción de la lesión isquémica, permitiendo así ajustar la representación para dar prioridad a esta región. Dicha función de pérdida se codifica como restricciones y mecanismos de ponderación, permitiendo una mayor relevancia de observación en una región en particular. La estrategia generativa emplea una red discriminadora PatchGAN y mapas de ponderación por tipo de tejido. Estos mapas integran información sobre la forma y el tamaño de las lesiones isquémicas, permitiendo resaltar patrones visuales en las secuencias sintéticas. La red discriminadora PatchGAN se utiliza para evaluar la autenticidad de las imágenes generadas, operando a nivel de pequeños parches de la imagen, lo que facilita una evaluación más detallada y local de las características de la imagen generada. Esta metodología se enfoca en la preservación y precisión de los patrones visuales asociados a las lesiones isquémicas, asegurando que las zonas críticas para el diagnóstico reciban mayor atención durante el proceso de generación de las imágenes sintéticas. Esta estrategia se entrenó y validó con un conjunto de datos que incluye 98 secuencias diagnósticas emparejadas de NCCT, ADC, DWI y FLAIR. Las imágenes sintéticas se

⁵ Heran YANG et al. "Unpaired Brain MR-to-CT Synthesis Using a Structure-Constrained CycleGAN". in: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Ed. by Danail STOYANOV et al. Cham: Springer International Publishing, 2018, pp. 174–182.

⁶ Karim ARMANIOUS et al. "Unsupervised medical image translation using cycle-MedGAN". in: *2019 27th European signal processing conference (EUSIPCO)*. IEEE. 2019, pp. 1–5.

⁷ Alejandro GUTIERREZ et al. "Lesion-preserving unpaired image-to-image translation between MRI and CT from ischemic stroke patients". In: *Medical Imaging 2022: Computer-Aided Diagnosis*. Vol. 12033. SPIE. 2022, pp. 322–328.

evaluaron mediante métricas de coherencia estructural, apariencia, textura y modelos de aprendizaje profundo, tanto a nivel global como en la zona de la lesión.

1. FUNDAMENTOS Y TRABAJOS PREVIOS

1.1. ACV isquémico y las imágenes médicas para soportar el diagnóstico

El ACV sucede cuando el suministro de sangre a una parte del cerebro se detiene debido a un bloqueo o ruptura en una arteria cerebral. Dependiendo de la causa, el ACV se clasifica en hemorrágico o isquémico. El ACV isquémico es la forma más frecuente de ACV con una incidencia de aproximadamente el 87% del total de los casos ². Este es causado por el estrechamiento u obstrucción de los vasos sanguíneos cerebrales por depósitos de grasa o coágulos que se forman en la sangre, y que viajan hasta el cerebro. Una vez que el flujo sanguíneo se ve afectado, el daño a nivel celular es rápido e irreversible, siendo el tratamiento temprano clave para disminuir la mortalidad y evitar daños adicionales que puedan afectar la calidad de vida de los pacientes ¹.

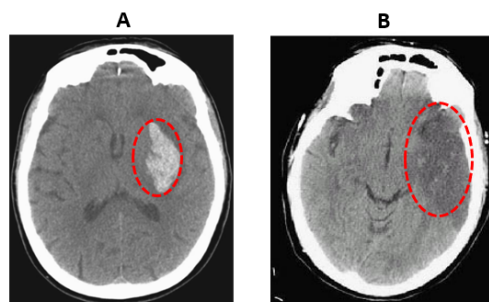
Los estudios imagenológicos son esenciales para el diagnóstico y el tratamiento del ACV isquémico, ya que permiten determinar la extensión y la localización de la lesión cerebral, así como identificar el tejido que aún puede ser recuperado. Los principales estudios imagenológicos que se utilizan son la TC (tomografía computarizada) y la RM (resonancia magnética), las cuales pueden realizarse con o sin contraste y con diferentes técnicas para evaluar el flujo sanguíneo, el metabolismo y la perfusión cerebral ⁸. La elección de la modalidad de imagen depende de la disponibilidad, el tiempo transcurrido desde el inicio de los síntomas, las características del paciente y los criterios clínicos ¹. En general, la TC es más rápida y accesible que la RM, y tiene una alta sensibilidad para detectar hemorragias cerebrales, las cuales contraindican el uso de trombolíticos. Por otro lado, la RM ofrece una mayor sensibilidad y especificidad que la TC, detectando

⁸ Scott RUDKIN et al. "Imaging of acute ischemic stroke". In: *Emergency radiology* 25 (2018), pp. 659–672.

lesiones isquémicas más tempranas y pequeñas. A continuación, se describen algunas de las modalidades más utilizadas para el diagnóstico del ACV isquémico.

Tomografía computarizada sin contraste (NCCT). La NCCT es la imagen diagnóstica más usada para evaluar a los pacientes con sospecha de ACV ⁸. Esta es rápida, económica y accesible, permitiendo distinguir entre ACV isquémico y hemorrágico. En particular, tiene una alta sensibilidad para detectar hemorragias intracraneales, donde las regiones afectadas se muestran con una mayor hiperdensidad en comparación al resto del parénquima cerebral. Sin embargo, la NCCT no es muy eficaz para detectar los ACV isquémicos agudos, sobre todo en las primeras horas desde el inicio de los síntomas. De hecho, un indicio de ACV isquémico es una zona más hipodensa, que muestra que hay menos sangre y más agua en el tejido cerebral dañado. Esta zona hipodensa suele ser muy tenue y difícil de ver en esta modalidad ⁹. (Ver Figura 1).

Figura 1. NCCT de dos pacientes diferentes: (A) ACV hemorrágico, donde la región afectada se muestra de forma hiperdensa. (B) ACV isquémico, donde la lesión se observa hipodensa.



Imágenes por resonancia magnética (RM). La RM es una técnica de imagen que ofrece información anatómica detallada del cerebro y otras estructuras. Particularmente, los estudios de RM ponderados por difusión son útiles en el diagnóstico del ACV isquémico

⁹ Romane GAURIAU et al. "Head CT deep learning model is highly accurate for early infarct estimation". In: *Scientific Reports* 13.1 (2023), p. 189.

agudo, permitiendo localizar el núcleo de la lesión, es decir, el tejido cerebral que ya está muerto, incluso en las primeras horas desde el inicio de los síntomas ⁸. Sin embargo, la RM también tiene algunas desventajas, como su alto costo, su menor disponibilidad, su mayor tiempo de adquisición y dificultad para monitorear movimientos bruscos del paciente. ¹⁰.

Difusión por resonancia magnética (DWI). Las imágenes DWI son la referencia principal para el diagnóstico imagenológico del ACV agudo, con una sensibilidad que oscila entre el 73% y el 92% durante las primeras 3 horas, y una sensibilidad cercana al 100% en las primeras 6 horas después del inicio de la isquemia ⁸. Además, la técnica DWI tiene la capacidad de identificar la isquemia aguda desde los primeros 30 minutos en que se presentan los síntomas, destacándose como el mejor método para cuantificar el volumen del tejido cerebral afectado por un ACV isquémico agudo ¹¹. En la Figura 2, se puede apreciar un ejemplo en el que los patrones visuales asociados a la isquemia son más evidentes en DWI en comparación con NCCT.

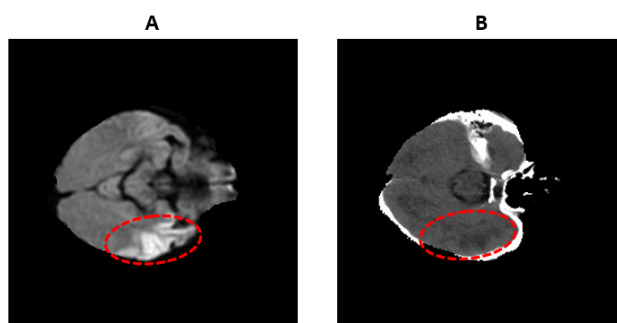
Un aspecto clave de la técnica DWI es el parámetro b , el cual indica la sensibilidad a la difusión de las moléculas de agua. Valores más altos de b , que oscilan entre 1500 y 5000 s/mm^2 , implican una mayor sensibilidad a los movimientos microscópicos de las moléculas de agua. Aunque los valores altos de b no siempre ofrecen ventajas diagnósticas adicionales en comparación con los valores estándar, pueden ser útiles en el diagnóstico de ACV hiperagudo o isquemia global ¹².

¹⁰ Linden PARKES et al. "An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI". in: *Neuroimage* 171 (2018), pp. 415–436.

¹¹ Thabele M LESLIE-MAZWI et al. "Endovascular stroke treatment outcomes after patient selection based on magnetic resonance imaging and clinical criteria". In: *JAMA neurology* 73.1 (2016), pp. 43–49.

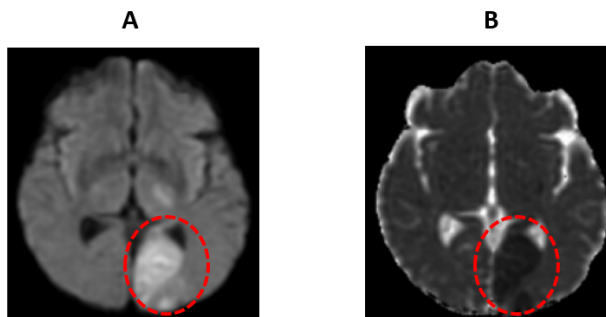
¹² Mutlu CIHANGIROGLU et al. "The utility of high b-value DWI in evaluation of ischemic stroke at 3 T". in: *European journal of radiology* 78.1 (2011), pp. 75–81.

Figura 2. DWI y NCCT de un mismo paciente: (A) DWI que muestra la lesión isquémica de forma hiperintensa. (B) NCCT que muestra un contraste bajo entre la lesión y el resto del tejido cerebral.



Coeficiente de difusión aparente (ADC). El ADC es un mapa que mide el movimiento de las moléculas de agua dentro de los tejidos. Este se calcula a partir de la atenuación de la señal en las imágenes DWI, utilizando diferentes valores de b (que reflejan la intensidad de la ponderación de la difusión). Usualmente, se emplean dos valores de b , típicamente $b = 0 \text{ s/mm}^2$ y $b = 1000 \text{ s/mm}^2$, para generar los mapas de ADC. La señal de las imágenes obtenidas con estos valores de b se utilizan para calcular el ADC mediante una relación logarítmica de las intensidades, que refleja la disminución de la señal causada por la difusión del agua ³. En una región afectada por un ACV isquémico, la restricción al movimiento de las moléculas de agua aumenta debido a la hinchazón celular, lo que resulta en valores de ADC significativamente bajos (regiones que se muestran hipointensas, ver Figura 3).

Figura 3. DWI y ADC de un mismo paciente: (A) DWI que muestra la lesión isquémica de forma hiperintensa. (B) ADC que muestra la lesión de manera hipointensa.



Recuperación de inversión atenuada por fluido (FLAIR). La secuencia FLAIR es crucial en el diagnóstico del ACV isquémico, ya que mejora la visualización de las lesiones al suprimir la señal del líquido cefalorraquídeo, permitiendo una detección más clara de las lesiones isquémicas ². Las imágenes FLAIR muestran las áreas isquémicas como hiperintensas. Además, el análisis conjunto de esta secuencia con DWI permite estimar la edad de la lesión, al tiempo que permite la diferenciación con otras patologías ^{13 2}. Particularmente estas secuencias FLAIR se complementan con el DWI, para en conjunto estimar el tiempo transcurrido del ACV. Este análisis en conjunto se denomina *DWI-FLAIR mismatch* y constituye una herramienta diagnóstica importante para determinar la ventana terapéutica en pacientes con ACV isquémico. Un mismatch DWI-FLAIR indica que una lesión es visible en DWI pero no en FLAIR (Figura 4), lo que sugiere que el ACV es reciente, generalmente dentro de las primeras 4.5 horas desde el inicio de los síntomas ¹³. Esta ventana es crítica para la administración de terapias de reperfusión, como la trombólisis intravenosa, que pueden ser efectivas para mejorar los resultados de los pacientes ¹³.

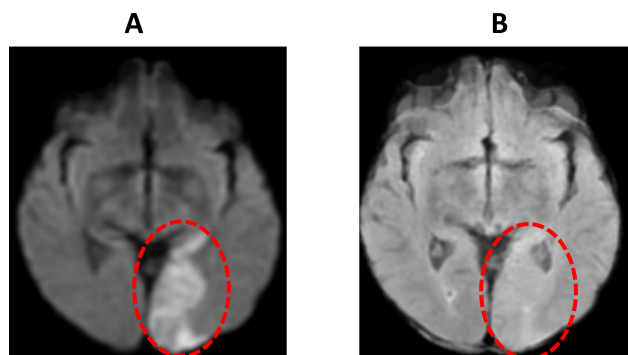
1.2. Estrategias de aprendizaje profundo para el aprendizaje generativo

La segmentación ha sido el principal método para caracterizar la lesión isquémica, ya que permite obtener una estimación del tejido afectado. Sin embargo, esta técnica no explota la información adicional que brindan las diferentes secuencias radiológicas ¹⁴. Otras técnicas se dedican a combinar los resultados radiológicos de varias modalidades de imagen para mejorar la representación de la lesión isquémica, lo que facilita a los

¹³ Götz THOMALLA et al. "DWI-FLAIR mismatch for the identification of patients with acute ischaemic stroke within 4-5 h of symptom onset (PRE-FLAIR): a multicentre observational study". In: *The Lancet Neurology* 10.11 (2011), pp. 978–986.

¹⁴ Hae Sol MOON et al. "Automated multimodal segmentation of acute ischemic stroke lesions on clinical MR images". In: *Magnetic Resonance Imaging* 92 (2022), pp. 45–57. DOI: [10.1016/j.mri.2022.06.001](https://doi.org/10.1016/j.mri.2022.06.001).

Figura 4. DWI y FLAIR de un mismo paciente: (A) DWI que muestra la lesión isquémica de forma hiperintensa. (B) FLAIR adquirido en las primeras 5 horas desde el inicio de los síntomas, donde se puede observar que aún no se presentan características hiperintensas asociadas a la lesión isquémica.



expertos el acceso a información relevante para decidir la intervención clínica del paciente ¹⁴. A continuación, se describen los principales mecanismos de estas tareas.

Los métodos generativos son una clase de algoritmos no supervisados, que buscan crear nuevas muestras a partir de los datos de entrenamiento ¹⁵. Estos algoritmos aprenden una representación de los datos, tratando de aproximar un espacio limitado que contiene a toda la población. Usando técnicas de aprendizaje profundo, se han desarrollado métodos generativos basados en el aprendizaje autosupervisado. Estos métodos se pueden dividir en dos grupos: los autocodificadores y las redes generativas adversarias (GAN). Los autocodificadores aprenden a comprimir y descomprimir los datos, ajustándose a una función de pérdida que puede contener tareas de reconstrucción de la información ¹⁶. Las GAN, en cambio, aprenden a generar un espacio latente que no depende de una distribución paramétrica ¹⁷. Un modelo GAN típico se entrena con un esquema especial,

¹⁵ Mehdi MIRZA and Simon OSINDER. “Conditional generative adversarial nets”. In: *arXiv preprint arXiv:1411.1784* (2014).

¹⁶ Walter Hugo Lopez PINAYA et al. “Autoencoders”. In: *Machine learning*. Elsevier, 2020, pp. 193–208.

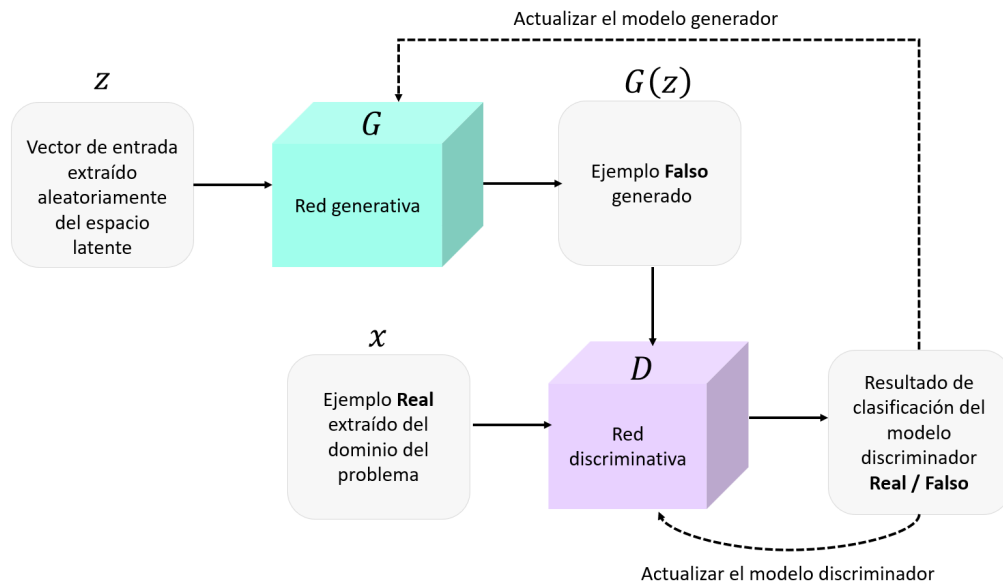
¹⁷ Ian GOODFELLOW et al. “Generative adversarial nets”. In: *Advances in neural information processing systems 27* (2014).

en el que una red que codifica y decodifica los datos aprende a crear muestras sintéticas (red generativa), mientras que otra red trata de distinguir las muestras reales de las falsas (red discriminativa). Estas redes se entrenan específicamente de acuerdo a la siguiente función de pérdida:

$$\mathcal{L}_{\text{GAN}}(G, D) = \min_G \max_D E_{x \sim P_{\text{data}}(x)} [\log D(x)] + E_{z \sim P_{\text{data}}(z)} [\log(1 - D(G(z)))] \quad (1)$$

Donde x corresponde a un valor real de entrada. La función $G(z)$ crea una muestra artificial a partir de un vector z que puede ser aleatorio o una imagen real, dependiendo del tipo de problema que se quiera resolver. Luego, la función D evalúa si la entrada es real o generada por $G(z)$.

Figura 5. Modelo generativo adversario (GAN). El generador (G) en verde, partiendo de un vector de ruido z , produce datos sintéticos intentando replicar los datos reales, mientras que el discriminador (D) en morado, evalúa tanto los datos reales x como los sintéticos $G(z)$, clasificándolos como verdaderos o falsos.

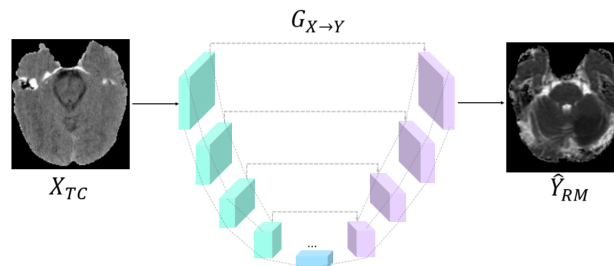


En la figura 5 se ilustra un modelo típico de GAN, en donde el generador (G) y el discriminador (D) son dos redes neuronales que se entrenan de forma conjunta y competitiva. El generador crea datos sintéticos a partir de un vector de ruido (z), tratando de imitar los

datos reales. Mientras tanto, el discriminador recibe datos reales (x) y sintéticos ($G(z)$), y trata de clasificarlos como verdaderos o falsos. El objetivo del generador es engañar al discriminador, produciendo datos tan realistas que el discriminador no pueda distinguirlos de los reales. A su vez, el objetivo del discriminador es identificar los datos que provienen del generador, detectando los datos falsos que el generador le envía. Ambas redes se entrenan de forma adversaria, es decir, compiten entre sí para mejorar su rendimiento. Este proceso de entrenamiento da lugar a un juego de suma cero en el que el equilibrio se alcanza cuando el generador produce datos que son indistinguibles de los datos reales, dificultando la capacidad de diferenciación del discriminador.

Traducción de dominio. La técnica de traducción o síntesis consiste en la reproducción o generación de una imagen artificial en un dominio diferente a la entrada. Por ejemplo, poder generar una imagen FLAIR a partir de una observación TC. particularmente este aprendizaje implica adaptar imágenes que tienen un conjunto de características (dominio X) a otro conjunto con diferentes características (dominio Y). Por ejemplo, obtener una secuencia de RM a partir de un estudio de TC, tal como se ilustra en la figura 6.

Figura 6. Síntesis de RM a partir de TC ($TC \rightarrow RM$). El modelo generativo $G_{X \rightarrow Y}$ sintetiza una imagen RM (\hat{Y}_{RM}) a partir de una imagen de TC (X_{TC}).



La estrategia principal para abordar el problema de traducción de imágenes entre domin-

ios se ha centrado en el uso de técnicas generativas profundas ¹⁸. Estas técnicas abarcan arquitecturas diseñadas tanto para datos emparejados, que logran mapear imágenes de un dominio X a un dominio Y manteniendo una correspondencia explícita entre ellas, como para datos no emparejados, capaces de aprender a establecer la correspondencia entre imágenes de un dominio X y un dominio Y sin necesidad de un emparejamiento explícito ¹⁹.

A continuación, profundizaremos en dos arquitecturas representativas del estado del arte: Pix2Pix y CycleGAN.

Pix2Pix. Es un modelo que se destaca por su capacidad para trabajar con datos emparejados, es decir, conjuntos de imágenes donde cada elemento en el dominio X tiene un correspondiente directo en el dominio Y ²⁰, lo que le permite aprender una función de mapeo específica para la tarea.

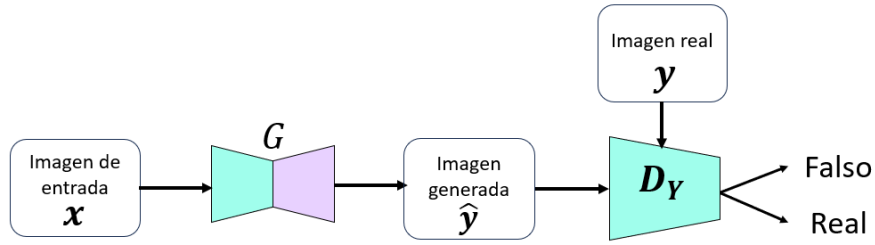
El entrenamiento de esta arquitectura se basa en una estrategia de optimización adversaria, como se ilustra en la figura 7. La función de objetivo es una combinación de la pérdida adversaria \mathcal{L}_{GAN} y la pérdida \mathcal{L}_{L_1} para controlar la reconstrucción de la imagen sintética (ver ecuación 2). La pérdida adversaria \mathcal{L}_{GAN} es crucial para garantizar la producción de imágenes realistas dentro del dominio objetivo, mientras que la pérdida \mathcal{L}_{L_1} se enfoca en preservar la integridad y exactitud de la imagen original perteneciente al dominio origen.

¹⁸ Gustavo GARZÓN et al. “A deep CT to MRI unpaired translation that preserve ischemic stroke lesions”. In: *2022 44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2022, pp. 2708–2711. DOI: [10.1109/EMBC48229.2022.9871154](https://doi.org/10.1109/EMBC48229.2022.9871154).

¹⁹ Jun-Yan ZHU et al. “Unpaired image-to-image translation using cycle-consistent adversarial networks”. In: *Proceedings of the IEEE international conference on computer vision*. 2017, pp. 2223–2232.

²⁰ Phillip ISOLA et al. “Image-to-image translation with conditional adversarial networks”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 1125–1134.

Figura 7. Modelo Pix2Pix.



$$\mathcal{L} = \mathcal{L}_{\text{GAN}}(G, D, X, Y) + \lambda \mathcal{L}_{L_1}(G) \quad (2)$$

La pérdida adversaria L_{GAN} se define como en la ecuación (1), mientras la pérdida \mathcal{L}_{L_1} se expresa como:

$$\mathcal{L}_{L_1}(G) = E_{x,y}[\|y - G(x)\|_1] \quad (3)$$

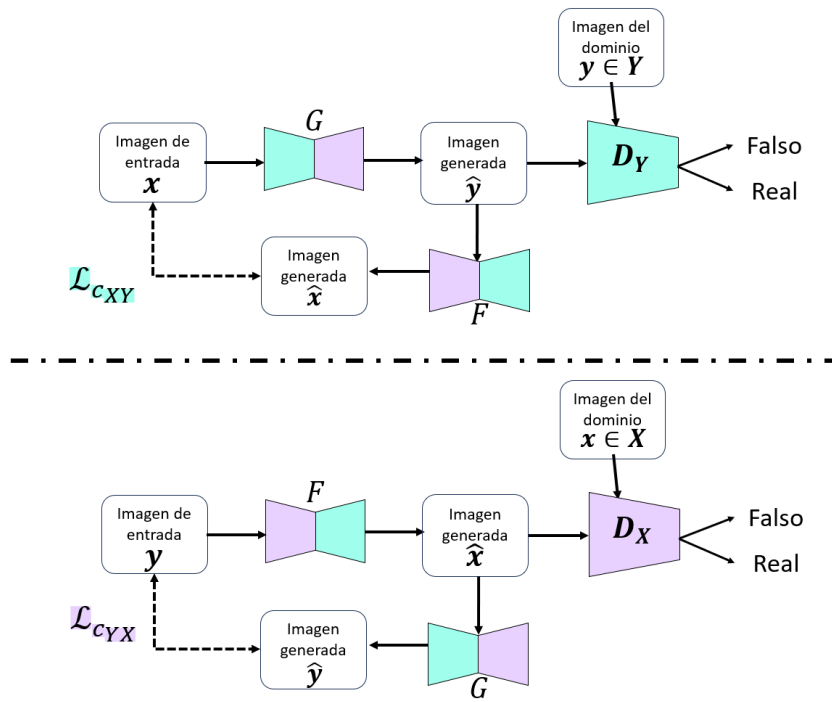
En este caso, $G(x)$ es la imagen producida por el generador a partir de la imagen de entrada x , y la pérdida busca minimizar la diferencia absoluta entre los píxeles de esta imagen generada y los de la imagen objetivo y . A su vez, λ es un hiperparámetro para equilibrar el proceso.

CycleGAN. Esta arquitectura ha sido ampliamente utilizada para hallar la correspondencia entre imágenes de un dominio X y un dominio Y sin que tengan que estar emparejadas ¹⁹. Este modelo involucra dos generadores: G y F . El generador G trata de pasar del dominio X a Y , sin embargo, esta estimación unidireccional podría producir resultados no deseados, como imágenes sintéticas deformadas o anatómicamente incorrectas. Por eso, una función generadora F soluciona este problema pasando del dominio Y al X , y una función discriminadora se encarga de diferenciar entre muestras reales y sintéticas (ver Figura 8). Con el fin de garantizar estas condiciones, se utiliza una pérdida de consistencia cíclica que asegura que las imágenes generadas sean muy similares a

las del espacio de origen de la proyección, es decir:

$$\mathcal{L}_{\text{cyc}}(G, F) = E_{X \sim P_{\text{data}}(x)} [\|F(G(x)) - x\|_1] + E_{Y \sim P_{\text{data}}(y)} [\|G(F(y)) - y\|_1] \quad (4)$$

Figura 8. Arquitectura de modelo CycleGAN que utiliza dos mapeos generativos, $G : X \rightarrow Y'$ (sección superior) y $F : Y \rightarrow X'$ (sección inferior), para transformar entre dominios no emparejados X y Y . G y F son complementados por discriminadores D_Y y D_X que diferencian entre imágenes reales y generadas



1.3. Mecanismos generativos y síntesis de imágenes en ACV

En la literatura se encuentran múltiples trabajos que abordan el problema de la traducción de dominio entre diferentes modalidades de neuroimágenes, empleando arquitecturas generativas. La metodología más utilizada consiste en ajustar una arquitectura de aprendizaje profundo de tipo autocodificador, a partir de conjuntos de datos no emparejados. Trabajos como el propuesto por Wolterink et al. diseñaron un modelo CycleGAN

para sintetizar imágenes cerebrales de TC a partir de imágenes cerebrales de RM ²¹. Yang et al. estimaron TC sintéticas a partir de RM, utilizando una arquitectura CycleGAN que restringe la consistencia estructural a partir de una función de costo, estimada desde un descriptor de vecindad independiente de la modalidad (MIND) ²². Este trabajo tiene la ventaja de no requerir delineaciones hechas por expertos para lograr obtener una consistencia estructural. También, Prokopenko et al. presentaron un método de síntesis de secuencias RM, a partir de TC que usa una estrategia basada en CycleGAN con una función de pérdida perceptual y una capa convolucional coordinada para mejorar la calidad de las imágenes sintéticas ²³. No obstante, esta técnica requiere un procesamiento previo y posterior para mejorar la calidad de la imagen. Armanious et al. idearon un modelo Cycle-MedGAN para transformar imágenes PET en imágenes TC con énfasis en la reducción de las diferencias de textura y percepción entre las imágenes originales y las sintéticas ⁶. En otro estudio posterior, Armanious et al. presentaron un método que combina un generador CasNet, un extractor VGG-19 y un discriminador PatchGAN para llevar a cabo la traducción de PET a TC, además de corregir los defectos debido al movimiento del paciente durante la adquisición de secuencias RM y eliminar el ruido de secuencias

²¹ Jelmer M WOLTERINK et al. “Deep MR to CT synthesis using unpaired data”. In: *Simulation and Synthesis in Medical Imaging: Second International Workshop, SASHIMI 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 10, 2017, Proceedings 2*. Springer. 2017, pp. 14–23.

²² Heran YANG et al. “Unpaired brain MR-to-CT synthesis using a structure-constrained CycleGAN”. in: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*. Springer. 2018, pp. 174–182.

²³ Denis PROKOPENKO et al. “Unpaired synthetic image generation in radiology using gans”. In: *Artificial Intelligence in Radiation Therapy: First International Workshop, AIRT 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Proceedings 1*. Springer. 2019, pp. 94–101.

PET ²⁴. Estos enfoques permiten aprender las características texturales más relevantes entre dos dominios de imágenes médicas, pero no garantizan la preservación de la estructura anatómica precisa entre los dominios. Esto puede provocar inconsistencias o distorsiones en las imágenes sintéticas, que afecten su calidad y utilidad clínica.

Específicamente en el contexto de ACV isquémico, Rubin y Mazdak diseñaron un método para sintetizar imágenes DWI a partir de mapas de CTP usando redes generativas adversarias condicionales (cGAN) ²⁵. Sin embargo, este trabajo tiene limitaciones en la evaluación cualitativa y cuantitativa de los datos sintéticos y validación clínica del método propuesto. En publicaciones más recientes, Garzon et al. idearon una estrategia que permite obtener secuencias sintéticas de DWI a partir de TC, enfocándose en la preservación de la estructura cerebral y las regiones de lesión isquémica ¹⁸. Este método utiliza en su arquitectura concatenaciones de módulos U-Net (U-blocks), y además, obtienen un espacio topológico y geométrico donde las secuencias sintéticas de DWI se proyectan como puntos embebidos en el espacio, y asumen que su posición se relaciona con la presencia de características de lesión isquémica. A pesar de ello, esta estrategia no tiene en cuenta la información semántica de las lesiones. Es decir, esta arquitectura no garantiza que las lesiones se mantengan en la misma ubicación, tamaño y forma después de la traducción. Por su parte, Gutiérrez et al. presentaron una estrategia de traducción de TC a imágenes de resonancia magnética FLAIR, basada también en un modelo CycleGAN, haciendo dos contribuciones importantes: la primera es la adición de una pérdida de consistencia de gradiente, y la segunda es la integración de módulos de atención (máscaras de lesiones generadas por otra red) en el proceso de traducción ⁷. Estos módulos permiten guiar

²⁴ Karim ARMANIOUS et al. "MedGAN: Medical image translation using GANs". In: *Computerized medical imaging and graphics* 79 (2020), p. 101684.

²⁵ Jonathan RUBIN and S Mazdak ABULNAGA. "CT-To-MR conditional generative adversarial networks for ischemic stroke lesion segmentation". In: *2019 IEEE International Conference on Healthcare Informatics (ICHI)*. IEEE. 2019, pp. 1–7.

el proceso de traducción asignando una mayor importancia a las regiones afectadas por lesiones isquémicas. No obstante, un problema que persiste en la traducción es el manejo de las lesiones de gran tamaño o de bajo contraste con el tejido normal, lo que produce errores en la representación de las intensidades propias de cada modalidad.

2. PROBLEMA DE INVESTIGACIÓN

El ACV es la segunda causa de mortalidad en el mundo, reportando una cifra de más de 6.5 millones de fallecimientos en el 2019 ²⁶. Particularmente en Colombia se reportaron 16.090 decesos en el 2018 ²⁷. Estas cifras alarmantes se deben principalmente al tipo más común de ACV: el isquémico, el cual representa aproximadamente un 87% del total de los casos ². El diagnóstico y tratamiento oportuno del ACV son factores cruciales para reducir el riesgo de muerte y la discapacidad de los pacientes. Para ello, se utilizan diferentes estudios de imagenología que permiten visualizar las lesiones cerebrales y distinguir entre ACV isquémico y hemorrágico, así como estimar la extensión y localización de la lesión ⁸. No obstante, estos estudios presentan desafíos en términos de disponibilidad, costo, tiempo y sensibilidad. Por ejemplo, la TC es un estudio ágil y económico, pero tiene una baja sensibilidad para detectar cambios en la fase aguda del ACV isquémico. Por otro lado, la RM proporciona imágenes más detalladas y sensibles, pero es un estudio costoso, prolongado y escaso. Estas limitaciones dificultan la valoración y el tratamiento adecuado de esta patología, ya que a menudo no se dispone de todas las modalidades que permiten identificar las lesiones isquémicas.

Ante este problema, se han propuesto algoritmos de aprendizaje profundo para estimar las imágenes diagnósticas faltantes a partir de las imágenes ya adquiridas. Estas estrategias se basan en métodos generativos que aprenden a traducir entre diferentes modalidades de imagenología, como TC y RM. Sin embargo, estas estrategias no modelan de

²⁶ Benjamin Stark VALERY FEIGIN Christopher J.L. Murray. “Global, regional, and national burden of stroke and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019”. In: *The Lancet Neurology* 20.10 (2021), pp. 795–820. DOI: [doi.org/10.1016/S1474-4422\(21\)00252-0](https://doi.org/10.1016/S1474-4422(21)00252-0).

²⁷ Ministerio de SALUD Y PROTECCIÓN SOCIAL. *Conozca cómo prevenir los accidentes cerebrovasculares*. <https://acortar.link/dHN0w9>. (Accessed on 09/13/2023). Oct. 2020.

manera explícita la información visual de la lesión, generando inconsistencias en las imágenes sintéticas como la incoherencia anatómica de las lesiones. El uso de estas técnicas puede ser clave para mejorar la preservación de las lesiones al enfatizar las regiones relevantes de la imagen y penalizar los errores en las regiones cerebrales afectadas por el ACV.

De esta manera, surge la siguiente pregunta de investigación:

¿Cómo implementar una estrategia generativa para resaltar los patrones visuales de las lesiones de accidente cerebrovascular isquémico durante la traducción entre secuencias radiológicas?

3. OBJETIVOS

3.1. Objetivo general

Implementar una estrategia generativa para la traducción entre secuencias radiológicas que resalte los patrones visuales de las lesiones de accidente cerebrovascular isquémico.

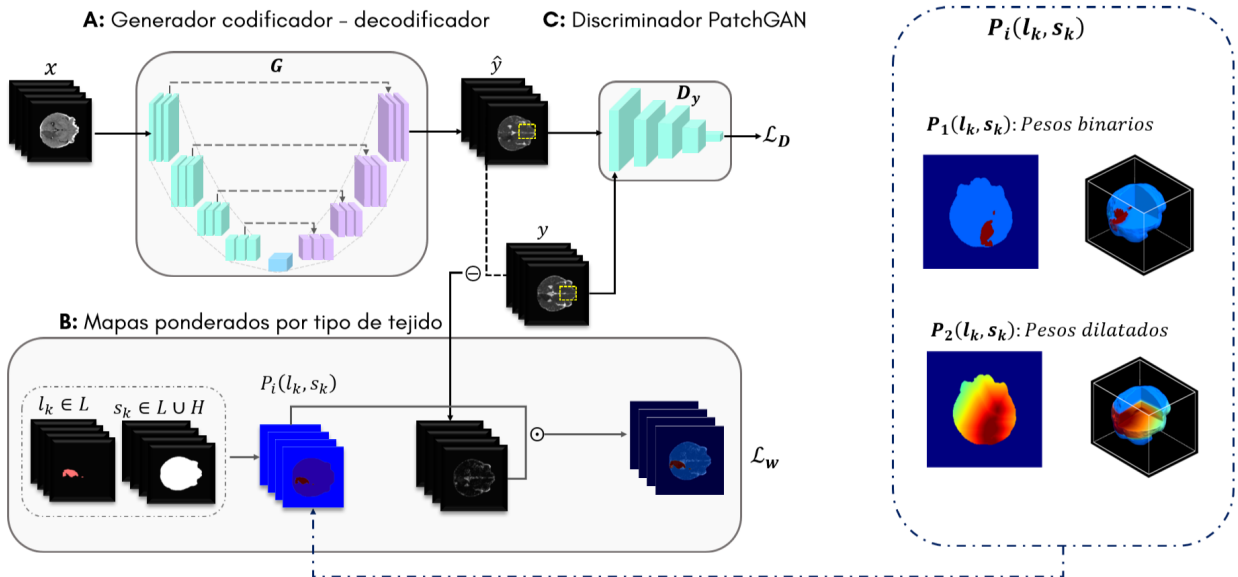
3.2. Objetivos específicos

- Procesar un conjunto de datos con al menos dos modalidades radiológicas de pacientes con ACV isquémico que incluyan anotaciones de las lesiones.
- Implementar un modelo generativo para la traducción entre secuencias radiológicas.
- Implementar una estrategia que incorpore la información de la lesión en la traducción entre secuencias radiológicas.
- Validar la estrategia implementada usando métricas de coherencia estructural, textural y realizando comparaciones detalladas en las regiones anotadas con lesión.

4. MÉTODO PROPUESTO

En este trabajo se implementó un método de traducción emparejada para secuencias diagnósticas de ACV, diseñado para preservar los patrones visuales de las lesiones en ambos dominios (ver Figura 9). Este método consta de dos redes convolucionales: un generador que predice los patrones tisulares más probables en el dominio de destino a partir de una imagen del dominio de origen. Por otra parte, el discriminador que se encarga de diferenciar entre imágenes reales y generadas. Para ajustar esta arquitectura, se incluyeron mapas ponderados por tipo de tejido, los cuales refuerzan el aprendizaje sobre el tejido hipoperfundido, mejorando la detección y caracterización de las lesiones. A continuación se presentan cada uno de los componentes del enfoque desarrollado.

Figura 9. Esquema del enfoque propuesto. Se entrena una arquitectura generativa (A) con imágenes de dos modalidades, incorporando mapas de pesos de clase (B) - binarios y dilatados - y utilizando un discriminador (C) para evaluar la calidad de las imágenes generadas.



4.1. Generador: una arquitectura codificador-decodificador

Para extraer los patrones tisulares relacionados con las lesiones isquémicas, y realizar una estimación precisa de la modalidad faltante, se desarrolló un generador G que sigue una arquitectura U-Net (ver Figura 9A). Esta arquitectura está compuesta por un codificador que comprime progresivamente las secuencias de entrada mediante bloques convolucionales y operaciones de pooling, los cuales condensan las características imagenológicas en descriptores de baja dimensionalidad. Posteriormente, esta representación sirve de entrada para una red decodificadora que utiliza bloques convolucionales y capas de convolución transpuesta para estimar los patrones estructurales y tisulares del dominio de referencia. Durante el proceso de decodificación, conexiones de salto enriquecen la información en cada nivel convolucional, mitigando la pérdida de detalles finos ocasionada por las operaciones de contracción en el codificador. Adicionalmente, estas conexiones de salto ayudan a contrarrestar el desvanecimiento del gradiente, un desafío técnico común en el entrenamiento de redes profundas.

Con el objetivo de reforzar la traducción de las lesiones isquémicas, propusimos una función de pérdida que toma en cuenta la delineación de la lesión isquémica e implementa pesos de clase binarios (p_1) o dilatados (p_2).

$$\mathcal{L}_w = \frac{1}{T} \sum_{i=1}^N p_m(l_k, s_k) |y_i - \hat{y}_i| \quad (5)$$

La pérdida del generador $\mathcal{L}_G = D(\hat{y}) + \lambda \mathcal{L}_w$ considera el término del discriminador $D(\hat{y})$, y lo suma a un bloque regularizador que incluye λ y la función de pérdida de pesos de clase \mathcal{L}_w . Particularmente, $p_m(l_k, s_k)$ es un mapa que mantiene la información espacial de la lesión y el tejido sano circundante, designado por s_k . Este mapa proporciona un contexto para evaluar los píxeles, asignando un incremento en la ponderación a los que pertenecen a la lesión, representados por l_k . Al hacerlo, se intensifica la penalización en la función de pérdida ajustando la diferencia absoluta entre el valor del píxel real y_k y el generado \hat{y}_k .

Esta mayor ponderación significa que las discrepancias en estas áreas críticas afectan de manera más significativa el cálculo de la pérdida, reforzando la precisión en la detección y representación de las lesiones isquémicas.

4.2. Discriminador PatchGAN

Con el fin de analizar elementos de pequeña escala en la evaluación de la autenticidad de las lesiones de ACV isquémico, se implementó un discriminador PatchGAN. Este discriminador es una variante especializada, que a diferencia de los discriminadores convencionales que analizan la imagen completa, este se enfoca en pequeños parches, clasificando cada uno por separado. Luego de procesar los n parches, el discriminador calcula un promedio de las clasificaciones individuales de cada parche para obtener una medida global de autenticidad para la imagen completa. Esto ha mostrado que en algunos trabajos del estado del arte, permite capturar detalles locales con mayor precisión y mejora la eficiencia ²⁰. Formalmente, la pérdida del discriminador se define como:

$$\mathcal{L}_D = \frac{1}{2} \left(\mathbb{E}_{(x,y) \sim p_{data}(x,y)} [\log(D(x, y))] + \mathbb{E}_{(x,\hat{y}) \sim p_{fake}(x,\hat{y})} [\log(1 - D(x, \hat{y}))] \right) \quad (6)$$

Aquí, \mathbb{E} representa la esperanza matemática o valor esperado. $(x, y) \sim p_{data}(x, y)$ indica que el par (x, y) (imagen de entrada x y su correspondiente imagen real y) se toma de la distribución de datos reales p_{data} . De manera similar $(x, \hat{y}) \sim p_{fake}(x, \hat{y})$, indica que el par (x, \hat{y}) (imagen de entrada x y su correspondiente imagen sintética \hat{y}) se toma de la distribución de datos generados p_{fake} . En ese sentido, el primer término es la probabilidad de que el discriminador clasifique correctamente una imagen real (x, y) como real. Y el segundo término, es la probabilidad de que el par $(x, \hat{y}) = (x, G(x))$ sea clasificado como falso. El objetivo del discriminador es minimizar esta pérdida para mejorar su capacidad de distinguir entre imágenes reales y generadas.

4.3. Optimización de la generación por tipo de tejido

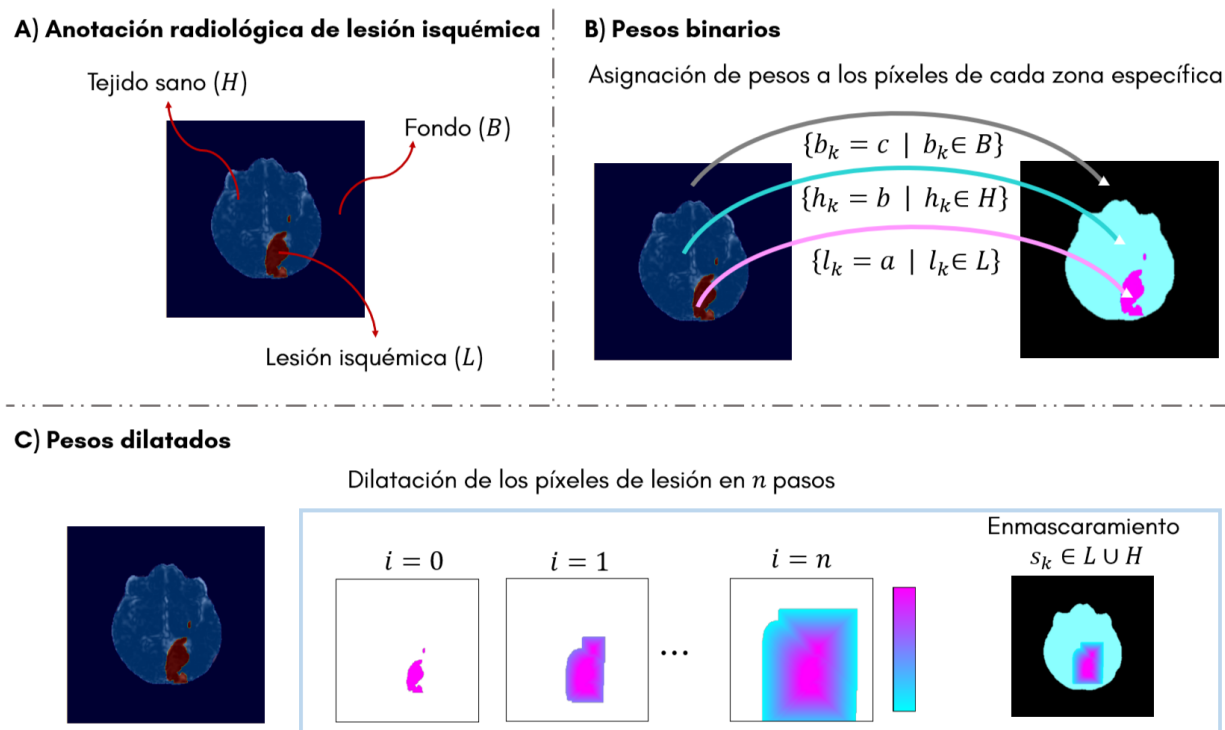
Para mejorar el proceso de traducción, en este trabajo decidimos incorporar explícitamente la delineación de la lesión por parte del experto clínico dentro de la función de pérdida (ver Figura 9B). Más precisamente, proponemos dos posibles esquemas de ponderación (pesos de clase binarios y dilatados) que operan sobre los píxeles de la lesión $l_k \in L$, los píxeles de tejido sano $h_k \in H$, y los píxeles de fondo $b_k \in B$. También consideramos una categoría adicional, a saber, los píxeles del cerebro s_k , correspondientes a la unión de los píxeles de lesión y de tejido sano $s_k \in L \cup H$. Este enfoque asegura que el modelo priorice las características clínicamente relevantes, mejorando la calidad y aplicabilidad clínica de las imágenes sintéticas.

Sea T el número total de píxeles en una imagen t_i , donde $T = L+H+B$ de tal forma que si la imagen t_i tiene una lesión, entonces $|L| > 0$. Además, sea $\{a, b, c\} \in \mathbb{R}$ hiperparámetros que corresponden a píxeles de lesión, tejido sano y fondo respectivamente. Los siguientes esquemas reciben tanto los píxeles de la lesión (l_k) como los píxeles del cerebro (s_k) como parámetros, y asumen la existencia de lesiones isquémicas en cualquier imagen t_i .

4.3.1. Pesos de clase binarios. Sea $p_1(l_k, s_k)$ una función que asigna el valor más alto a a los píxeles de la lesión isquémica $\{l_k = a \mid l_k \in L\}$ y el valor más bajo c a los píxeles de fondo $\{b_k = c \mid b_k \in B\}$, mientras que asigna un valor intermedio b a los píxeles de tejido sano $\{h_k = b \mid h_k \in H\}$. (Ver Figura 10B)

4.3.2. Pesos de clase dilatados. Considere $p_2(l_k, s_k)$ como una función que aplica una operación de dilatación binaria morfológica a la región de píxeles de la lesión a través de p pasos de dilatación. Los píxeles que pertenecen a cada paso de dilatación contendrán p valores intermedios únicos entre el valor más alto a y el valor de tejido sano b . Además, para establecer límites anatómicos, los pasos de dilatación morfológica fueron

Figura 10. Aplicación de métodos de ponderación. En la sección A de la figura se muestra un corte axial extraído de un volumen de un estudio radiológico, el cual incluye marcaciones específicas de la lesión isquémica. En las secciones B y C, se ilustran los dos métodos de ponderación aplicados sobre la misma imagen: el método binario en la Sección A y el método dilatado en la Sección B.



enmascarados por los píxeles de corte de cráneo $\{s_k \in L \cup H\}$ para esa imagen, de modo que los píxeles fuera de esa máscara pertenecen al valor de fondo c . (Ver Figura 10C)

La pérdida total del modelo resulta de la combinación de la pérdida del generador y la pérdida del discriminador, ya que ambos componentes se optimizan simultáneamente durante el proceso de entrenamiento. Por lo tanto, la expresión de la pérdida total del modelo es la siguiente:

$$\mathcal{L}_{total} = \mathcal{L}_G + \mathcal{L}_D \quad (7)$$

5. DISEÑO EXPERIMENTAL

5.1. Datos

El conjunto de datos empleado para este trabajo incluye registros de 98 pacientes, a quienes se les realizó una NCCT al momento de su ingreso al centro médico. Dentro de las siguientes 24 horas, se llevaron a cabo estudios de RM en las modalidades de DWI, ADC y FLAIR. Dos neurorradiólogos con más de cinco años de experiencia, efectuaron anotaciones de las lesiones isquémicas en las imágenes ADC.

5.2. Preprocesamiento

Con el fin de aislar el tejido cerebral, se le removió el cráneo a todas las secuencias diagnósticas utilizando el software SynthStrip ²⁸. Particularmente, para las imágenes NCCT fue necesario llevar a cabo una reconstrucción aproximada de un T1 MP-RAGE antes para obtener una remoción del cráneo. Este proceso convierte las NCCT en imágenes que resaltan la estructura cerebral, facilitando una segmentación precisa del tejido cerebral ²⁹. Seguidamente, se realizó un registro rígido de las imágenes NCCT, DWI y FLAIR hacia las imágenes ADC utilizando Elastix ³⁰. Este procedimiento implica ajustar las imágenes para que coincidan en posición y orientación mediante transformaciones que incluyen solo rotaciones y traslaciones, sin alterar la escala o aplicar deformaciones.

²⁸ Andrew HOOPES et al. "SynthStrip: Skull-stripping for any brain image". In: *NeuroImage* 260 (2022), p. 119474.

²⁹ Juan E IGLESIAS et al. "SynthSR: A public AI tool to turn heterogeneous clinical brain scans into high-resolution T1-weighted images for 3D morphometry". In: *Science advances* 9.5 (2023), eadd3607.

³⁰ Denis P SHAMONIN et al. "Fast parallel image registration on CPU and GPU for diagnostic classification of Alzheimer's disease". In: *Frontiers in neuroinformatics* 7 (2014), p. 50.

Se empleó una interpolación de B-spline de tercer orden para optimizar la precisión de los valores de intensidad de los píxeles durante la transformación. Este paso asegura que todas las modalidades estén correctamente alineadas con las localizaciones de las anotaciones de las lesiones, facilitando así análisis comparativos y de seguimiento más precisos. Para mejorar la visualización de las imágenes NCCT se aplicó un ventaneo con un centro de ventana en 40 Hounsfield units (HU) y un ancho de ventana de 80 HU. Posteriormente, se normalizaron las intensidades de cada vóxel entre 0 y 1 para luego ser guardadas como PNG.

5.3. Configuración de la arquitectura

El generador fue construido con cuatro niveles de codificación y decodificación. Los niveles de codificación comienzan con 64 filtros y se duplican en cada nivel subsiguiente hasta alcanzar los 512 filtros en el cuello de botella. Cada capa de codificación utiliza convoluciones con un kernel de tamaño 4×4 , seguidas de normalización de lotes y activaciones *LeakyReLU*. Durante la fase de decodificación, se emplean convoluciones transpuestas con mismo tamaño de kernel, seguidas de normalización de lotes y activaciones *ReLU*. La imagen final es producida mediante una convolución transpuesta y una activación *Tanh*. Los pesos para los mapas de clase se fijaron en 1 para los píxeles de fondo, 3 para los píxeles de cerebro y 10 para los píxeles de lesión. Respecto al discriminador PatchGAN, consta de tres capas convolucionales, cada capa utiliza un kernel de tamaño 4×4 con activaciones *LeakyReLU* y normalización de lotes, el tamaño del parche utilizado en esta evaluación es de 70×70 píxeles. Las imágenes de entrada y sus máscaras se preprocesan mediante redimensionamiento y normalización antes de ser procesadas por la red. Las imágenes de entrada utilizadas tienen dimensiones de 256×256 píxeles. En cuanto al entrenamiento, se empleó el optimizador Adam con una tasa de aprendizaje inicial de 0.0002 y $\beta_1 = 0.5$. La tasa de aprendizaje se ajusta mediante un decaimiento lineal a lo largo de las últimas 100 épocas, después de mantener una tasa constante

durante las primeras 100 épocas de entrenamiento, totalizando 200 épocas.

5.4. Validación

Para la validación de la arquitectura desarrollada se crearon múltiples versiones del conjunto de datos con dos objetivos principales, evaluar la capacidad de nuestro método para traducir entre diferentes secuencias de imágenes y preservar los detalles de las lesiones isquémicas. La Tabla 1 muestra las características específicas de cada par de modalidades de imagen analizadas. Debido a problemas en la calidad de 11 imágenes FLAIR, fue necesario descartarlas.

Tabla 1. Distribución de pacientes por pares de modalidades de imagen. Las filas en la tabla 1 que incluyen un asterisco (*) indican los pares de modalidades que involucran la modalidad FLAIR.

Síntesis	Entrenamiento	Evaluación
NCCT → FLAIR*	52	35
DWI → FLAIR*	52	35
ADC → FLAIR*	52	35
NCCT → ADC	60	38
NCCT → DWI	60	38

5.4.1. Métricas de evaluación. Para evaluar la calidad de las imágenes sintéticas generadas en este estudio, se consideraron métricas estandarizadas y basadas en modelos pre-entrenados que permiten una comparación objetiva y cuantitativa de los resultados. Estas métricas proporcionan una visión amplia de la fidelidad de las imágenes, abarcando desde la similitud estructural hasta la percepción visual. A continuación, se detallan las métricas utilizadas:

Proporción máxima de señal-ruido (PSNR). La proporción máxima de señal-ruido (PSNR, por sus siglas en inglés) es una métrica ampliamente utilizada para evaluar la calidad de imágenes sintéticas en comparación con las originales. El PSNR se calcula a

partir del error cuadrático medio (MSE), que es una medida fundamental del error promedio entre la imagen original $I(x, y)$ y la imagen generada $I'(x, y)$. El MSE se define como:

$$MSE = \frac{1}{MN} \sum_{y=1}^M \sum_{x=1}^N [I(x, y) - I'(x, y)]^2 \quad (8)$$

donde M y N son las dimensiones de la imagen, representando el número de filas y columnas, respectivamente. Utilizando el MSE, el PSNR se calcula mediante la fórmula:

$$PSNR = 10 \cdot \log_{10} \left(\frac{\max_I^2}{MSE} \right) \quad (9)$$

Aquí, \max_I representa el valor máximo posible de un píxel en la imagen. Un PSNR alto indica que la diferencia entre la imagen reconstruida y la original es menor, lo cual es indicativo de una mayor calidad en la traducción de la imagen.

Índice de similitud estructural (SSIM). El índice de similitud estructural (SSIM, por sus siglas en inglés) es una métrica que evalúa la calidad de las imágenes no solo en términos de error promedio, como el MSE, sino también considerando cambios en la estructura, luminosidad y contraste. El SSIM se basa en la premisa de que la percepción humana de las imágenes es altamente sensible a las variaciones estructurales. En ese sentido, esta métrica se define mediante la fórmula 10:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (10)$$

Aquí, μ_x y μ_y son los promedios de las intensidades de los píxeles, σ_x y σ_y son las varianzas, y σ_{xy} es la covarianza entre las imágenes x e y . Las constantes C_1 y C_2 se añaden para mantener la estabilidad numérica. Un valor alto de SSIM (cercano a 1) indica una alta similitud entre la imagen generada y la original.

Índice de similitud estructural multiescala (MS-SSIM). El índice de similitud estructural multiescala (MS-SSIM, por sus siglas en inglés) extiende la idea del SSIM para evaluar la calidad de la imagen en múltiples escalas o resoluciones. Esto es particularmente útil para reconocer diferencias en detalles finos y estructuras más amplias que podrían perderse en una sola escala. La fórmula del MS-SSIM es:

$$MS-SSIM = [l_M(x, y)]^{\alpha_M} \cdot \prod_{j=1}^M [c_j(x, y)]^{\beta_j} [s_j(x, y)]^{\gamma_j} \quad (11)$$

En esta fórmula, $l_M(x, y)$ representa la similitud en luminancia en la escala más alta, mientras que $c_j(x, y)$ y $s_j(x, y)$ miden la similitud en contraste y estructura en la escala j . Los exponentes α_M , β_j , y γ_j controlan la importancia relativa de cada componente. Un valor más alto de MS-SSIM indica una mejor preservación de la calidad a través de múltiples escalas.

Distancia de inicio de Frechet (FID). La métrica FID (por sus siglas en inglés) se basa en la distancia entre las distribuciones de características obtenidas mediante la red Inception-v3. La FID evalúa cómo las características extraídas de un conjunto de imágenes generadas difieren estadísticamente de las de un conjunto de imágenes reales ³¹. Un valor bajo de FID indica que las imágenes generadas son estadísticamente similares a las reales, lo que sugiere una alta calidad en la generación.

La FID se calcula mediante la diferencia de las medias y la traza de la diferencia de covarianzas entre las características extraídas de las imágenes reales y las generadas, como se muestra en la fórmula 12:

$$FID = \|\mu_r - \mu_g\|^2 + Tr(\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{1/2}) \quad (12)$$

³¹ Martin HEUSEL et al. "Gans trained by a two time-scale update rule converge to a local nash equilibrium". In: *Advances in neural information processing systems* 30 (2017).

donde μ_r, Σ_r y μ_g, Σ_g son las medias y covarianzas de las características de las imágenes reales y generadas, respectivamente.

Similitud perceptiva aprendida entre parches de imágenes (LPIPS). La métrica LPIPS (por sus siglas en inglés) mide las diferencias perceptuales entre dos imágenes utilizando características profundas extraídas por redes neuronales pre-entrenadas. A diferencia de métricas que comparan imágenes a nivel de píxeles, LPIPS evalúa cómo las diferencias son percibidas visualmente por humanos, utilizando redes como VGG, AlexNet o SqueezeNet ³². Un valor bajo de LPIPS indica una alta similitud perceptual entre las imágenes comparadas, sugiriendo una calidad visual superior.

5.4.2. Prueba estadística. Para evaluar la significancia estadística de los resultados obtenidos en la comparación de los diferentes métodos de síntesis implementados, se utilizó la Prueba de Mann-Whitney U. Esta prueba permite comparar las métricas de dos muestras independientes y estimar si las diferencias observadas no son producto del azar. Esta prueba se aplicó para analizar si existían diferencias significativas entre cada método de síntesis utilizando pesos binarios y dilatados con respecto al método que no incorpora dichos pesos, así como entre los dos métodos (binario y dilatado). El valor ρ se definió en 0.05 para determinar la significancia estadística de los resultados. Un $\rho < 0.05$ indica que hay una diferencia significativa entre los grupos comparados, mientras que un $\rho > 0.05$ sugiere que no hay evidencia suficiente para afirmar que existen diferencias significativas.

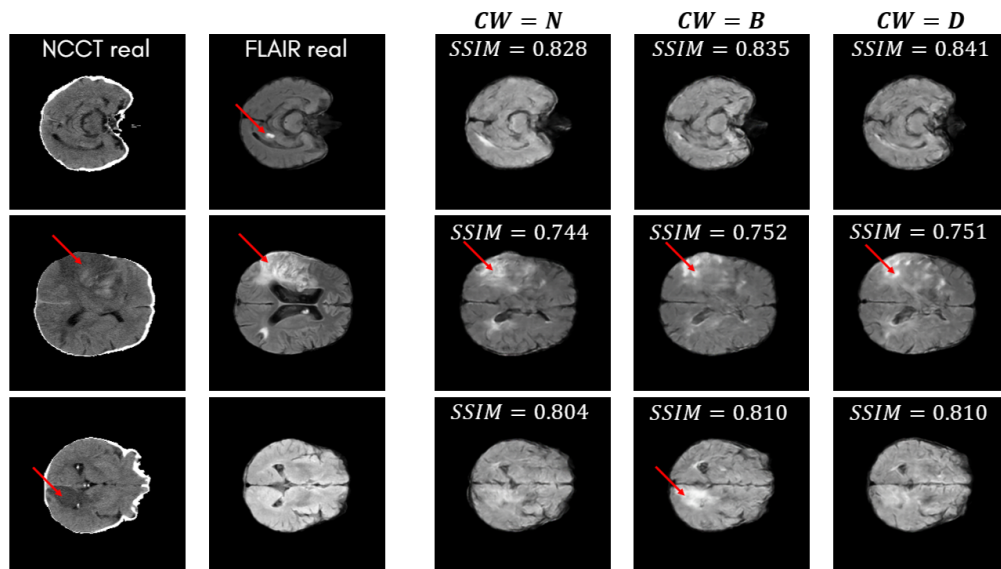
³² Richard ZHANG et al. "The unreasonable effectiveness of deep features as a perceptual metric". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 586–595.

6. EVALUACIÓN Y RESULTADOS

6.1. Secuencias FLAIR sintéticas como soporte en el análisis de DWI-FLAIR mismatch

En una primera evaluación, se decidió hacer síntesis de imágenes FLAIR para soportar el procedimiento de diferenciación junto con las imágenes DWI. La síntesis de estas imágenes se realizó desde las secuencias NCCT, DWI y mapas ADC, respectivamente. En general, las secuencias FLAIR sintéticas preservaron la información de la lesión isquémica y mostraron un rendimiento preciso en las métricas estructurales.

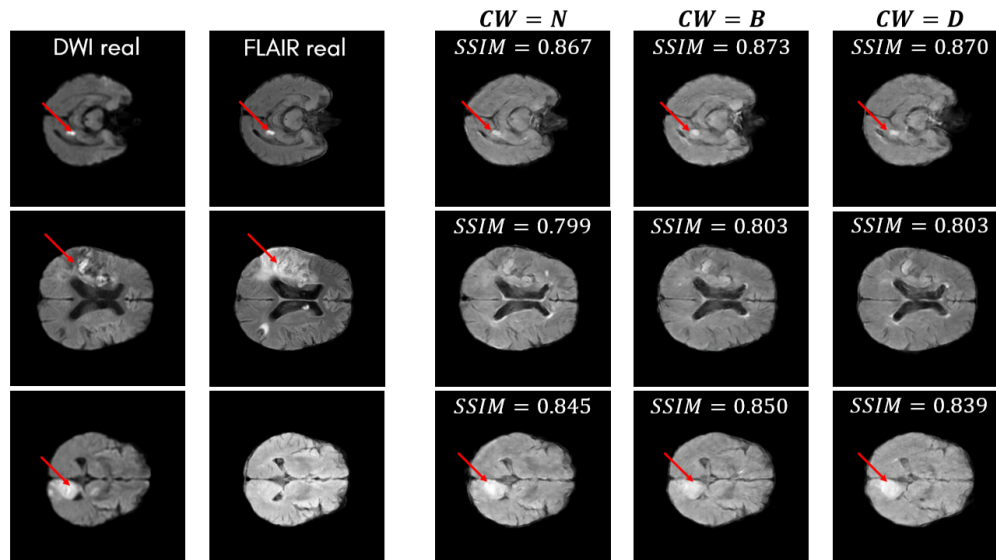
Figura 11. Resultados cualitativos de las secuencias FLAIR sintéticas a partir de NCCT. Cada fila corresponde a lesiones isquémicas diminutas (0.796 mL - TDIS: 3 h, TDIS: Tiempo desde inicio de los síntomas), pequeñas (25.5 mL - TDIS: despertar) y medianas (56.6 mL - TDIS: 5 h), respectivamente. Se denomina *CW* a los pesos de clase (Class Weights, en inglés).



En la Figura 11, las dos primeras filas presentan imágenes FLAIR reales con una flecha roja que indica la lesión. Las secuencias FLAIR sintetizadas desde NCCT sin pesos de clase (*CW* = *N*) en la tercera columna tienen menor contraste y claridad en la lesión,

además de algunos errores en la estructura del cerebro. Mientras que, al emplear pesos de clase binarios ($CW = B$) y dilatados ($CW = D$) en la cuarta y quinta columna, se observa una preservación adecuada de la lesión y de la estructura anatómica. Por otro lado, en la tercera fila, la FLAIR real no muestra características asociadas a la lesión, pero las imágenes sintetizadas con pesos binarios y dilatados logran preservar esta lesión, específicamente al utilizar pesos binarios, se logra una visibilidad con un contraste más acertado.

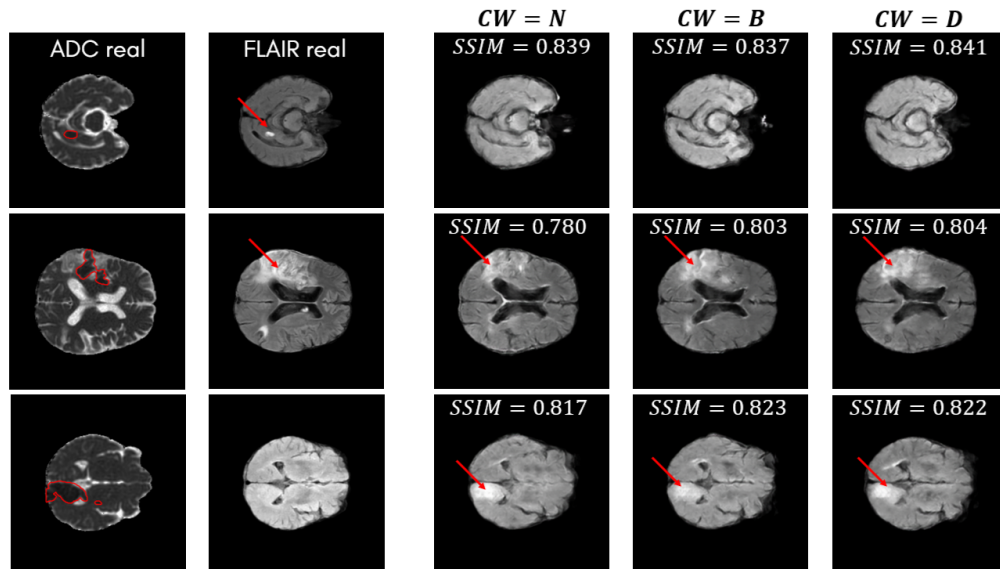
Figura 12. Resultados cualitativos de las secuencias FLAIR sintéticas a partir de DWI. Cada fila corresponde a lesiones isquémicas diminutas (0.796 mL - TDIS: 3 h, TDIS: Tiempo desde inicio de los síntomas), pequeñas (25.5 mL - TDIS: despertar) y medianas (56.6 mL - TDIS: 5 h), respectivamente.



En la Figura 12 se ilustran algunos ejemplos de las imágenes FLAIR sintetizadas desde DWI utilizando los tres métodos de ponderación, considerados en el método desarrollado ($CW = N$, $CW = B$, $CW = D$) muestran una representación adecuada de la lesión. Sin embargo, al usar pesos de clase binarios ($CW = B$) y dilatados ($CW = D$), la estructura anatómica del cerebro se aproxima más a la imagen real. En particular, en la tercera fila, la lesión es visible en DWI pero no en FLAIR, y las imágenes sintetizadas logran preservar la visibilidad y el detalle de la lesión. Este caso específico es un candidato ideal

para estudiar el DWI-FLAIR mismatch, ya que la imagen fue adquirida en menos de 6 horas desde el inicio de los síntomas. Esta discrepancia entre las imágenes DWI y FLAIR puede proporcionar información valiosa para analizar la evolución temprana del ACV y optimizar el tratamiento.

Figura 13. Resultados cualitativos de las secuencias FLAIR sintéticas a partir de mapas ADC. Cada fila corresponde a lesiones isquémicas diminutas (0.796 mL - TDIS: 3 h, TDIS: Tiempo desde inicio de los síntomas), pequeñas (25.5 mL - TDIS: despertar) y medianas (56.6 mL - TDIS: 5 h), respectivamente.



En la Figura 13 se ilustran las imágenes FLAIR sintéticas obtenidas a partir de mapas ADC. Particularmente en esta imagen se muestran los mapas ADC que contienen la delineación experta de la lesión isquémica. Cabe destacar que los radiólogos marcaron sobre estos mapas, lo cual se usa como referencia en nuestro enfoque para generar las secuencias sintéticas. En la primera fila, que corresponde a una lesión diminuta, ninguno de los tres métodos ($CW = N$, $CW = B$, $CW = D$) logró preservar la lesión de manera efectiva en las imágenes FLAIR sintetizadas. Mientras que, en la segunda fila, los tres métodos muestran la lesión, pero al usar pesos de clase binarios ($CW = B$) y dilatados ($CW = D$), la representación de la lesión es más aproximada a la imagen FLAIR original. Por otro lado, en la tercera fila, los tres métodos también logran mantener la lesión visible,

pero la estructura anatómica del cerebro no es tan precisa sin usar pesos de clase (CW = N). En cambio, al usar pesos dilatados (CW = D), la estructura anatómica se aproxima más a la imagen real.

Tabla 2. Métricas de evaluación de calidad estructural y perceptual de imágenes FLAIR sintéticas a partir de NCCT, DWI y mapas ADC. La columna de pesos de clase (CW) corresponde a ninguno (N), binario (B) y dilatado (D).

Síntesis	CW	SSIM \uparrow	PSNR \uparrow	MS-SSIM \uparrow	FID \downarrow	LPIPS (0,1) \downarrow
NCCT \rightarrow FLAIR	N	0.871 \pm 0.023	22.309 \pm 0.973	0.865 \pm 0.031	0.017	0.205 \pm 0.064
	B	0.874\pm0.023	22.498\pm0.978	0.870 \pm 0.032	0.011	0.206 \pm 0.064
	D	0.873 \pm 0.023	22.444 \pm 1.028	0.873\pm0.031	0.013	0.205 \pm 0.064
DWI \rightarrow FLAIR	N	0.895 \pm 0.024	23.622 \pm 1.050	0.897 \pm 0.031	0.016	0.204\pm0.064
	B	0.902\pm0.025	24.003\pm1.130	0.905\pm0.030	0.022	0.204 \pm 0.065
	D	0.897 \pm 0.025	23.819 \pm 1.127	0.900 \pm 0.023	0.016	0.206 \pm 0.064
ADC \rightarrow FLAIR	N	0.884 \pm 0.021	23.218 \pm 0.911	0.884 \pm 0.031	0.006	0.206 \pm 0.064
	B	0.890 \pm 0.022	23.613 \pm 1.032	0.891 \pm 0.032	0.01	0.206 \pm 0.064
	D	0.891\pm0.022	23.629\pm1.079	0.892\pm0.032	0.011	0.205\pm0.064

En cuanto a los resultados cuantitativos, la Tabla 2 muestra el rendimiento global durante el estudio retrospectivo descrito en este trabajo al obtener imágenes FLAIR sintéticas a partir de mapas NCCT, DWI y ADC. La síntesis de NCCT a FLAIR alcanzó una mejora de SSIM/PSNR del 0.3%/0.19 con pesos de clase binarios y un aumento en MS-SSIM del 0.8% con pesos dilatados.

Además, la síntesis de DWI a FLAIR logró un aumento en SSIM/PSNR/MS-SSIM del 0.7%/0.38/0.8% al implementar pesos de clase binarios. De manera similar, la síntesis de ADC a FLAIR produjo un aumento en SSIM/PSNR/MS-SSIM del 0.7%/0.41/0.8% al utilizar pesos dilatados. En cuanto a las métricas perceptuales, la síntesis de NCCT a FLAIR obtuvo una disminución en FID/LPIPS de 0.006/0.001 al implementar pesos binarios y dilatados respectivamente. Del mismo modo, la síntesis de DWI a FLAIR mostró una disminución en FID de 0.006 al usar pesos dilatados, mientras que LPIPS no mostró

cambios. Asimismo, los mapas de ADC a FLAIR no mostraron cambios en FID, mientras que se obtuvo una disminución en LPIPS de 0.001 al utilizar pesos dilatados.

Tabla 3. Métricas de evaluación de calidad estructural y perceptual de imágenes FLAIR sintéticas, delimitadas por la región de la lesión isquémica, a partir de NCCT, DWI y mapas ADC. La columna de pesos de clase (CW) corresponde a ninguno (N), binario (B) y dilatado (D).

Síntesis	CW	SSIM \uparrow	PSNR \uparrow	MS-SSIM \uparrow	FID \downarrow	LPIPS (0,1) \downarrow
NCCT \rightarrow FLAIR	N	0.474 \pm 0.060	14.549 \pm 1.639	0.468 \pm 0.055	0.143	0.384 \pm 0.048
	B	0.482\pm0.059	15.440\pm1.328	0.484\pm0.056	0.068	0.378 \pm 0.050
	D	0.477 \pm 0.060	15.108 \pm 1.519	0.469 \pm 0.068	0.110	0.377\pm0.051
DWI \rightarrow FLAIR	N	0.553 \pm 0.0062	16.348 \pm 1.466	0.584 \pm 0.047	0.107	0.346 \pm 0.042
	B	0.556 \pm 0.038	16.495\pm1.562	0.585 \pm 0.073	0.103	0.344\pm0.042
	D	0.563\pm0.046	16.300 \pm 1.615	0.604\pm0.047	0.127	0.345 \pm 0.042
ADC \rightarrow FLAIR	N	0.505 \pm 0.063	16.103 \pm 1.764	0.513 \pm 0.064	0.039	0.364 \pm 0.051
	B	0.531\pm0.054	17.211\pm1.326	0.545\pm0.055	0.069	0.357\pm0.048
	D	0.526 \pm 0.057	16.918 \pm 1.449	0.539 \pm 0.055	0.051	0.359 \pm 0.044

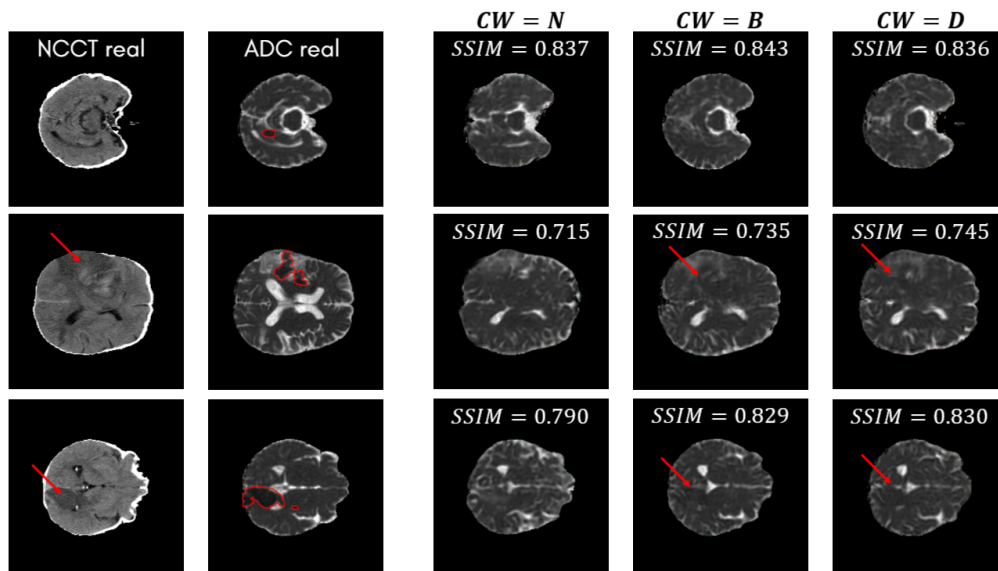
En particular, la Tabla 3 muestra el desempeño sobre nuestro conjunto de datos al obtener imágenes FLAIR sintéticas a partir de mapas NCCT, DWI y ADC con el mecanismo de las lesiones, para regiones delimitadas por la lesión isquémica. Por ejemplo, en comparación con no usar nuestro método de recuperación de lesiones, la conversión de NCCT a FLAIR obtuvo un aumento de SSIM/PSNR/MS-SSIM de 0.8%/0.9/1.6% usando pesos de clase binarios. Asimismo, la conversión de DWI a FLAIR obtuvo una ganancia de SSIM/MS-SSIM de 1%/2% usando pesos de clase dilatados, y un aumento de PSNR de 0.15 usando pesos de clase binarios. De manera similar, los mapas ADC a FLAIR lograron una ganancia de SSIM/PSNR/MS-SSIM de 2.6%/1.11/3.2% implementando pesos de clase binarios. En cuanto a las métricas perceptuales, solo la conversión de NCCT a FLAIR obtuvo una disminución de FID de 0.075 usando pesos de clase binarios. En lo que respecta a LPIPS, la conversión de NCCT a FLAIR obtuvo una pérdida de 0.007 usando pesos de clase dilatados, la conversión de DWI a FLAIR obtuvo una disminución de

0.002, y los mapas ADC a FLAIR lograron una pérdida de 0.007 usando pesos de clase binarios.

6.2. DWI y mapas ADC sintéticos

En este trabajo también se consideró la generación de secuencias DWI y mapas ADC, permitiendo así completar el flujo de trabajo clínico para análisis de pacientes con sospecha de ACV. Así, se procedió a generar secuencias DWI y mapas ADC sintéticos a partir de imágenes NCCT debido a la mayor sensibilidad de DWI y ADC para detectar cambios en la fase aguda del ACV isquémico. Esta síntesis se realizó utilizando las configuraciones de pesos de clases ya mencionadas. En general, las secuencias DWI y ADC sintéticas preservaron detalles estructurales críticos y mostraron un rendimiento preciso en las métricas de calidad visual y estructural.

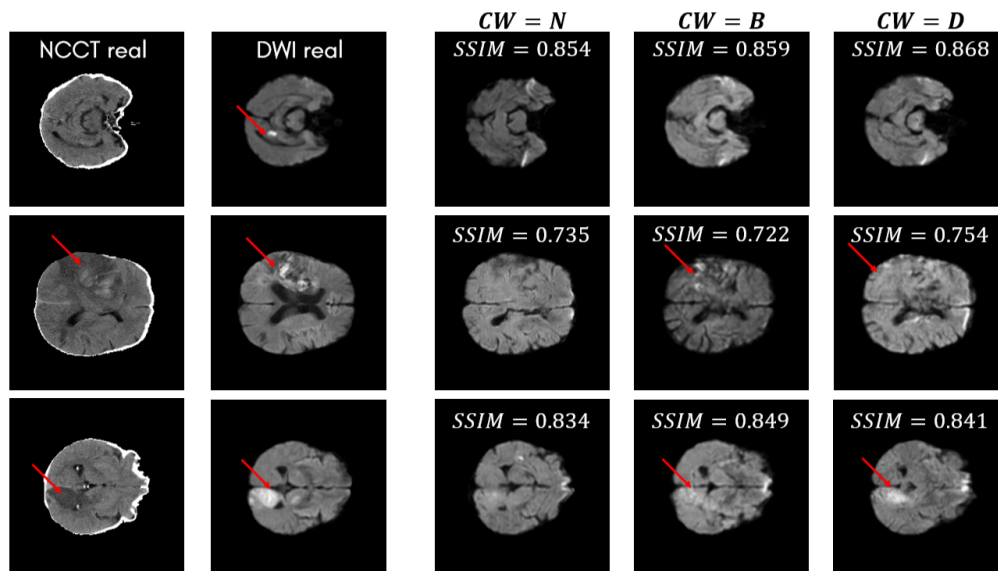
Figura 14. Resultados cualitativos de mapas ADC sintéticos a partir de NCCT. Cada fila corresponde a lesiones isquémicas diminutas (0.796 mL - TDIS: 3 h, TDIS: Tiempo desde inicio de los síntomas), pequeñas (25.5 mL - TDIS: despertar) y medianas (56.6 mL - TDIS: 5 h), respectivamente.



En el caso de la Figura 14, se ilustran los mapas ADC sintéticos obtenidos a partir de

secuencias NCCT. En la primera fila, que corresponde a una lesión diminuta, ninguno de los tres métodos ($CW = N$, $CW = B$, $CW = D$) logra preservar la lesión de manera efectiva en las imágenes ADC sintéticas. No obstante, la estructura anatómica del cerebro es más precisa y cercana a la imagen real cuando se utilizan pesos de clase binarios ($CW = B$) y dilatados ($CW = D$). Por otro lado, en las filas 2 y 3, la lesión se pierde por completo al no usar pesos de clase ($CW = N$). Sin embargo, al emplear pesos de clase binarios ($CW = B$) y dilatados ($CW = D$), la lesión se logra mantener, relativamente visible. Además, en todas las filas, la estructura anatómica es más precisa y cercana a la real al utilizar estos pesos de clase en comparación con no utilizarlos.

Figura 15. Resultados cualitativos de secuencias DWI sintéticas a partir de NCCT. Cada fila corresponde a lesiones isquémicas diminutas (0.796 mL - TDIS: 3 h, TDIS: Tiempo desde inicio de los síntomas), pequeñas (25.5 mL - TDIS: despertar) y medianas (56.6 mL - TDIS: 5 h), respectivamente.



En la Figura 15, se ilustran las imágenes DWI sintéticas obtenidas a partir de las secuencias NCCT. En la primera fila, la lesión no se preserva con ninguno de los tres métodos. Sin embargo, al no usar pesos de clase ($CW = N$), la imagen sintética se aleja considerablemente de la imagen real, mientras que al utilizar pesos de clase binarios ($CW = B$) y dilatados ($CW = D$), las imágenes sintéticas preservan mejor la estructura anatómica.

En la segunda fila, aunque los tres métodos parecen aproximarse a la lesión, al no usar pesos de clase ($CW = N$), la región afectada por isquemia presenta intensidades incorrectas, mostrándose hipointensa en lugar de hiperintensa como debería ser. En cambio, al utilizar pesos de clase binarios ($CW = B$) y dilatados ($CW = D$), la traducción de la lesión, aunque no es exacta, muestra la intensidad correcta. Por último, en la tercera fila, al utilizar pesos de clase ($CW = B$ y $CW = D$), se logra una traducción adecuada de la lesión. Específicamente, al usar pesos dilatados ($CW = D$), el contraste de la lesión es más acertado.

Tabla 4. Evaluación de la calidad estructural y perceptual de mapas ADC/imágenes DWI sintéticas a partir de NCCT. La columna de pesos de clase (CW) corresponde a ninguno (N), binario (B) y dilatado (D).

Síntesis	CW	SSIM \uparrow	PSNR \uparrow	MS-SSIM \uparrow	FID \downarrow	LPIPS (0,1) \downarrow
NCCT \rightarrow ADC	N	0.844 \pm 0.016	21.596 \pm 0.770	0.838 \pm 0.023	0.06	0.203 \pm 0.062
	B	0.867 \pm 0.020	22.722 \pm 0.978	0.876 \pm 0.025	0.082	0.200 \pm 0.064
	D	0.869\pm0.019	22.906\pm0.994	0.880\pm0.026	0.076	0.199\pm0.064
NCCT \rightarrow DWI	N	0.853 \pm 0.038	20.700 \pm 2.199	0.849 \pm 0.058	0.069	0.199 \pm 0.061
	B	0.874 \pm 0.024	21.524 \pm 1.363	0.864 \pm 0.029	0.045	0.194\pm0.061
	D	0.878\pm0.022	21.619\pm1.267	0.864\pm0.028	0.038	0.194 \pm 0.062

Respecto a los resultados cuantitativos, la Tabla 4 muestra el desempeño global al obtener mapas de ADC sintéticos y secuencias DWI. La síntesis desde NCCT hacia mapas de ADC obtuvo una ganancia de SSIM/PSNR/MS-SSIM de 2.5%/1.31/4.2% con CW dilatado. En cuanto a la síntesis desde NCCT a DWI, se obtuvo un aumento de SSIM/PSNR de 2.5%/0.91 al implementar CW dilatado, y una ganancia de MS-SSIM de 1.5% al usar CW binario. En lo que respecta a las métricas perceptuales, la síntesis de NCCT a mapas de ADC no mostró cambios en la métrica FID, mientras que LPIPS obtuvo una disminución de 0.004 al implementar CW dilatado. De manera similar, la síntesis de NCCT a DWI obtuvo una disminución de FID/LPIPS de 0.031/0.005 al usar CW dilatado y binario, respectivamente.

Tabla 5. Evaluación de la calidad estructural y perceptual de mapas ADC/imágenes DWI sintéticas, delimitadas por la región de la lesión isquémica, a partir de NCCT. La columna de pesos de clase (CW) corresponde a ninguno (N), binario (B) y dilatado (D).

Síntesis	CW	SSIM \uparrow	PSNR \uparrow	MS-SSIM \uparrow	FID \downarrow	LPIPS (0,1) \downarrow
NCCT \rightarrow ADC	N	0.514 \pm 0.088	15.567 \pm 1.926	0.499 \pm 0.111	0.091	0.361 \pm 0.057
	B	0.581\pm0.084	16.914\pm1.972	0.583\pm0.103	0.149	0.329\pm0.054
	D	0.573 \pm 0.081	16.787 \pm 1.930	0.574 \pm 0.104	0.121	0.329 \pm 0.055
NCCT \rightarrow DWI	N	0.493 \pm 0.073	13.867 \pm 2.508	0.454 \pm 0.089	0.129	0.372 \pm 0.058
	B	0.546 \pm 0.090	14.141 \pm 1.887	0.501 \pm 0.083	0.079	0.344 \pm 0.060
	D	0.558\pm0.086	14.836\pm1.727	0.522\pm0.091	0.101	0.340\pm0.058

Específicamente, la Tabla 5 muestra el rendimiento sobre nuestro conjunto de datos al obtener mapas de ADC sintéticos y secuencias DWI, para regiones delimitadas por la lesión isquémica. Por ejemplo, la síntesis de NCCT a mapas de ADC logró una ganancia de SSIM/PSNR/MS-SSIM de 6.7%/1.35/8.4% al usar CW binario. En cuanto a la síntesis de NCCT a DWI, se obtuvo un aumento de SSIM/PSNR/MS-SSIM de 6.5%/0.97/6.8% al usar CW dilatado. En lo que respecta a las métricas perceptuales, los mapas de ADC sintéticos obtuvieron una disminución de LPIPS de 0.032 al usar CW binario, y no mostraron diferencias en FID. En cuanto a DWI sintético, se obtuvo una disminución de FID/LPIPS de 0.05/0.032 al usar CW binario y dilatado, respectivamente.

En el estudio sobre la significancia estadística de la estrategia para evidenciar el aporte de la ponderación por clase introducida en este trabajo, se utilizó la prueba de Mann-Whitney U para evaluar las diferencias en las medianas de las métricas de calidad bajo diferentes configuraciones de pesos de clase. Esta prueba reveló diferencias significativas al usar pesos binarios o dilatados, comparados con la ausencia de pesos ($\rho < 0.05$), indicando que las configuraciones con pesos influyen positivamente en los resultados. Además, se confirmó que no hay diferencias estadísticamente significativas entre las configuraciones de pesos binarios y dilatados ($\rho > 0.05$), sugiriendo que ambos tipos de ponderación por clase producen efectos similares en las imágenes sintéticas. Estos resultados confirman

el impacto significativo de la incorporación de pesos en el entrenamiento, mejorando la percepción de las imágenes generadas.

7. CONCLUSIONES Y TRABAJO FUTURO

La atención oportuna en la detección y tratamiento de lesiones por ACV es crucial para disminuir la afectación neurológica y mejorar el pronóstico de los pacientes. Las modalidades de imagen médica, como FLAIR, DWI y mapas ADC, son fundamentales para un diagnóstico preciso. Estas imágenes no solo permiten cuantificar el volumen de la lesión, sino que también son esenciales para establecer tratamientos y la prognosis del paciente. Sin embargo, estas secuencias no se consideran parte central de la atención inicial, según las guías clínicas, y además su adquisición no se hace de forma inmediata. Nuestro método propuesto permite la recuperación eficiente de patrones de lesiones observados en estas secuencias a partir de estudios NCCT, proporcionando una herramienta valiosa en escenarios clínicos donde el tiempo es crítico para la toma de decisiones médicas.

En este trabajo se desarrolló un método de síntesis emparejado para secuencias diagnósticas de ACV, diseñado para preservar los patrones visuales de las lesiones isquémicas. Este método constó de un generador para predecir los patrones tisulares más probables en el dominio de destino a partir de una imagen del dominio de origen, y un discriminador encargado de diferenciar entre imágenes reales y generadas. Para ajustar esta arquitectura, se incluyeron mapas ponderados por tipo de tejido, los cuales reforzaron el aprendizaje sobre la región hipoperfundida del cerebro, mejorando la detección y caracterización de las lesiones. Para el entrenamiento y validación de esta estrategia se utilizó un conjunto de datos con secuencias diagnósticas emparejadas de NCCT, DWI, FLAIR y mapas ADC. Las imágenes sintéticas se evaluaron mediante métricas de coherencia estructural, apariencia, textura y modelos de aprendizaje profundo, tanto a nivel global como en la zona de la lesión.

El método desarrollado logró recuperar aproximaciones de patrones de lesiones típicamente observados en secuencias FLAIR a partir de exploraciones NCCT, DWI y ADC,

lo que podría permitir evaluar el DWI-FLAIR mismatch. En el contexto de un esquema emparejado, nuestro método con énfasis en la lesión sintetizó lesiones isquémicas sobre imágenes FLAIR con pesos de clase binarios y dilatados, mostrando una mejora en la calidad de las imágenes sintéticas.

Asimismo, la generación de secuencias DWI y mapas ADC a partir de NCCT demostró ser efectiva en la detección temprana de cambios asociados al ACV isquémico. Nuestro enfoque permitió la preservación de detalles estructurales y la mejora de las métricas de calidad visual y estructural. Al implementar pesos de clase binarios y dilatados, se obtuvieron imágenes sintéticas de DWI y ADC con una precisión adecuada, destacándose en la recuperación de detalles importantes y en la fidelidad de las lesiones isquémicas.

Cualitativamente, el enfoque propuesto evidenció patrones radiológicos destacados asociados con signos de lesiones isquémicas. De hecho, en el contexto de la traducción de la región isquémica, nuestro método obtuvo secuencias FLAIR sintéticas con una calidad estructural y perceptual adecuada. Además, la estrategia propuesta llevó a cabo un proceso que preservó los patrones de lesiones isquémicas que ocurren en la etapa inicial al implementar pesos de clase binarios y dilatados. Además, se mantuvieron los detalles anatómicos y estructurales de los tejidos cerebrales, así como los patrones perceptuales, que son clave para obtener secuencias FLAIR sintéticas aceptables.

En cuanto a las secuencias DWI y mapas ADC sintéticos, se observó una adecuada preservación de las lesiones isquémicas y una representación anatómica precisa al utilizar pesos de clase binarios y dilatados. Las imágenes sintéticas de DWI y ADC mostraron mejoras significativas en la visibilidad y detalle de las lesiones, permitiendo una evaluación más precisa y detallada en comparación con métodos que no utilizan pesos de clase. Estos resultados cualitativos evidencian la efectividad del método propuesto en la síntesis de imágenes diagnósticas complementarias para el ACV isquémico.

El enfoque de énfasis en las lesiones demostró ser efectivo en la síntesis de imágenes, mostrando mejoras significativas en las métricas de calidad visual y estructural tanto para

secuencias FLAIR como DWI y mapas ADC. Específicamente, los pesos de clase binarios y dilatados permitieron una mejor recuperación de detalles en las lesiones isquémicas, obteniendo mejores puntuaciones en las evaluaciones comparadas con las técnicas que no utilizan mecanismos de ponderación. Este resultado sugiere que nuestra metodología no solo mejora la precisión visual de las imágenes sintéticas sino que también preserva la integridad estructural necesaria para la evaluación clínica. Sin embargo, nuestro enfoque tiene limitaciones, especialmente en casos con bajo contraste entre la lesión y el tejido sano, y en lesiones diminutas. Estos desafíos son atribuibles a las características intrínsecas de las secuencias NCCT durante las etapas aguda e hiperaguda del ACV isquémico, así como al desequilibrio de clases entre cortes afectados y sanos, que pueden afectar el aprendizaje del patrón de lesiones isquémicas.

Para el trabajo futuro, nos enfocaremos en mejorar la técnica de síntesis para lesiones de bajo contraste y tamaño diminuto. Se investigarán otros métodos para abordar el desequilibrio de clases y se explorará la aplicación del método en cohortes con un mayor número de pacientes para validar su eficacia en un contexto clínico. Además, se realizarán estudios para cuantificar el impacto del intervalo de tiempo entre las adquisiciones de NCCT y las secuencias FLAIR, DWI y mapas ADC, para entender mejor cómo las diferencias temporales afectan la calidad de las imágenes sintetizadas y la precisión de la detección de lesiones. Con el fin de complementar la evaluación cuantitativa de los patrones de las lesiones isquémicas sintéticas, se explorarán métricas como el error absoluto promedio (MAE), el error cuadrático promedio (MSE), kernel inception distance (KID), inception score (IS) y estudios de percepción de realismo visual.

BIBLIOGRAFÍA

ARMANIOUS, Karim et al. “MedGAN: Medical image translation using GANs”. In: *Computerized medical imaging and graphics* 79 (2020), p. 101684 (cit. on p. 27).

ARMANIOUS, Karim et al. “Unsupervised medical image translation using cycle-MedGAN”. In: *2019 27th European signal processing conference (EUSIPCO)*. IEEE. 2019, pp. 1–5 (cit. on pp. 13, 26).

CIHANGIROGLU, Mutlu et al. “The utility of high b-value DWI in evaluation of ischemic stroke at 3 T”. In: *European journal of radiology* 78.1 (2011), pp. 75–81 (cit. on p. 17).

FILOGRANA, Laura et al. “Diffusion weighted imaging (DWI) and apparent diffusion coefficient (ADC) values for detection of malignant vertebral bone marrow lesions.” In: *European Review for Medical & Pharmacological Sciences* 22.3 (2018) (cit. on pp. 12, 18).

GARZÓN, Gustavo et al. “A deep CT to MRI unpaired translation that preserve ischemic stroke lesions”. In: *2022 44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2022, pp. 2708–2711. DOI: [10.1109/EMBC48229.2022.9871154](https://doi.org/10.1109/EMBC48229.2022.9871154) (cit. on pp. 23, 27).

GAURIAU, Romane et al. “Head CT deep learning model is highly accurate for early infarct estimation”. In: *Scientific Reports* 13.1 (2023), p. 189 (cit. on p. 16).

GOODFELLOW, Ian et al. “Generative adversarial nets”. In: *Advances in neural information processing systems* 27 (2014) (cit. on p. 20).

GUTIERREZ, Alejandro et al. “Lesion-preserving unpaired image-to-image translation between MRI and CT from ischemic stroke patients”. In: *Medical Imaging 2022: Computer-Aided Diagnosis*. Vol. 12033. SPIE. 2022, pp. 322–328 (cit. on pp. 13, 27).

HEUSEL, Martin et al. “Gans trained by a two time-scale update rule converge to a local nash equilibrium”. In: *Advances in neural information processing systems* 30 (2017) (cit. on p. 41).

HOOPEES, Andrew et al. “SynthStrip: Skull-stripping for any brain image”. In: *NeuroImage* 260 (2022), p. 119474 (cit. on p. 37).

IGLESIAS, Juan E et al. “SynthSR: A public AI tool to turn heterogeneous clinical brain scans into high-resolution T1-weighted images for 3D morphometry”. In: *Science advances* 9.5 (2023), eadd3607 (cit. on p. 37).

ISOLA, Phillip et al. “Image-to-image translation with conditional adversarial networks”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 1125–1134 (cit. on pp. 23, 34).

LESLIE-MAZWI, Thabele M et al. “Endovascular stroke treatment outcomes after patient selection based on magnetic resonance imaging and clinical criteria”. In: *JAMA neurology* 73.1 (2016), pp. 43–49 (cit. on p. 17).

MAKKAT, Smitha et al. “Signs of acute stroke seen on fluid-attenuated inversion recovery MR imaging”. In: *American Journal of Roentgenology* 179.1 (2002), pp. 237–243 (cit. on p. 12).

MIRZA, Mehdi and OSINDERO, Simon. “Conditional generative adversarial nets”. In: *arXiv preprint arXiv:1411.1784* (2014) (cit. on p. 20).

MOON, Hae Sol et al. “Automated multimodal segmentation of acute ischemic stroke lesions on clinical MR images”. In: *Magnetic Resonance Imaging* 92 (2022), pp. 45–57. DOI: [10.1016/j.mri.2022.06.001](https://doi.org/10.1016/j.mri.2022.06.001) (cit. on pp. 19, 20).

PARDO TURRIAGO, Rodrigo. “Guía de práctica clínica para el diagnóstico, tratamiento y rehabilitación del episodio agudo de ataque cerebrovascular en población mayor de 18 años: de las recomendaciones a su implementación”. In: *Acta Neurológica Colombiana* 31.4 (2015), pp. 462–467 (cit. on pp. 11, 12, 15).

PARKES, Linden et al. “An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI”. In: *Neuroimage* 171 (2018), pp. 415–436 (cit. on p. 17).

PINAYA, Walter Hugo Lopez et al. “Autoencoders”. In: *Machine learning*. Elsevier, 2020, pp. 193–208 (cit. on p. 20).

PROKOPENKO, Denis et al. “Unpaired synthetic image generation in radiology using gans”. In: *Artificial Intelligence in Radiation Therapy: First International Workshop, AIRT 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Proceedings 1*. Springer. 2019, pp. 94–101 (cit. on p. 26).

RUBIN, Jonathan and ABULNAGA, S Mazdak. “CT-To-MR conditional generative adversarial networks for ischemic stroke lesion segmentation”. In: *2019 IEEE International Conference on Healthcare Informatics (ICHI)*. IEEE. 2019, pp. 1–7 (cit. on p. 27).

RUDKIN, Scott et al. “Imaging of acute ischemic stroke”. In: *Emergency radiology* 25 (2018), pp. 659–672 (cit. on pp. 15–17, 29).

SALUD Y PROTECCIÓN SOCIAL, Ministerio de. *Conozca cómo prevenir los accidentes cerebrovasculares*. <https://acortar.link/dHN0w9>. (Accessed on 09/13/2023). Oct. 2020 (cit. on p. 29).

SHAMONIN, Denis P et al. “Fast parallel image registration on CPU and GPU for diagnostic classification of Alzheimer’s disease”. In: *Frontiers in neuroinformatics* 7 (2014), p. 50 (cit. on p. 37).

THOMALLA, Götz et al. “DWI-FLAIR mismatch for the identification of patients with acute ischaemic stroke within 4·5 h of symptom onset (PRE-FLAIR): a multicentre observational study”. In: *The Lancet Neurology* 10.11 (2011), pp. 978–986 (cit. on p. 19).

VALERY FEIGIN Christopher J.L. Murray, Benjamin Stark. “Global, regional, and national burden of stroke and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019”. In: *The Lancet Neurology* 20.10 (2021), pp. 795–820. DOI: [doi.org/10.1016/S1474-4422\(21\)00252-0](https://doi.org/10.1016/S1474-4422(21)00252-0) (cit. on p. 29).

WOLTERINK, Jelmer M et al. “Deep MR to CT synthesis using unpaired data”. In: *Simulation and Synthesis in Medical Imaging: Second International Workshop, SASHIMI 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 10, 2017, Proceedings 2*. Springer. 2017, pp. 14–23 (cit. on p. 26).

XIANG, Wei et al. “FLAIR vascular hyperintensity combined with asymmetrical prominent veins in acute anterior circulation ischemic stroke: prediction of collateral circulation and clinical outcome”. In: *European Journal of Medical Research* 28.1 (2023), p. 446 (cit. on pp. 11, 12, 15, 19, 29).

YANG, Heran et al. “Unpaired Brain MR-to-CT Synthesis Using a Structure-Constrained CycleGAN”. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for*

Clinical Decision Support. Ed. by Danail STOYANOV et al. Cham: Springer International Publishing, 2018, pp. 174–182 (cit. on p. 13).

YANG, Heran et al. “Unpaired brain MR-to-CT synthesis using a structure-constrained CycleGAN”. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*. Springer. 2018, pp. 174–182 (cit. on p. 26).

ZHANG, Richard et al. “The unreasonable effectiveness of deep features as a perceptual metric”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 586–595 (cit. on p. 42).

ZHU, Jun-Yan et al. “Unpaired image-to-image translation using cycle-consistent adversarial networks”. In: *Proceedings of the IEEE international conference on computer vision*. 2017, pp. 2223–2232 (cit. on pp. 23, 24).