

**DESARROLLO DE UN MODELO COMPUTACIONAL PARA LA GENERACIÓN DE
IMÁGENES ARTIFICIALES POST-CONTRASTE DE MRI ORIENTADO A LA
SÍNTESIS DE REGIONES CON POSIBLES LESIONES TUMORALES EN TEJIDOS DE
MAMA**



JULIANA MONTOYA DUQUE

**INSTITUCIÓN UNIVERSITARIA PASCUAL BRAVO
FACULTAD DE INGENIERÍA
INGENIERÍA DE SOFTWARE
MEDELLÍN
2024**

**DESARROLLO DE UN MODELO COMPUTACIONAL PARA LA GENERACIÓN DE
IMÁGENES ARTIFICIALES POST-CONTRASTE DE MRI ORIENTADO A LA
SÍNTESIS DE REGIONES CON POSIBLES LESIONES TUMORALES EN TEJIDOS DE
MAMA**

JULIANA MONTOYA DUQUE

Trabajo de grado para optar al título de Ingeniería de Software

Asesor

**RUBEN DARIO FONNEGRA TARAZONA
PhD. en Ingeniería**

**INSTITUCIÓN UNIVERSITARIA PASCUAL BRAVO
FACULTAD DE INGENIERÍA
INGENIERÍA DE SOFTWARE
MEDELLÍN
2024**

Contenido

	Pág.
Introducción	10
1. Planteamiento del problema	11
1.1 Descripción	11
1.2 Formulación	12
2. Justificación	13
3. Objetivos	14
3.1 Objetivo general	14
3.2 Objetivos específicos	14
4. Marco teórico	15
5. Metodología	19
5.1 Tipo de proyecto	19
5.2 Método	19
5.3 Instrumentos de recolección de información	30
6. Resultados del proyecto	31
7. Conclusiones	37
8. Referencias bibliográficas	38

Lista de figuras

	Pág.
<i>Figura 1. Imágenes de ejemplo extraídas de la base de datos. La imagen de la izquierda corresponde con la precontraste, la de en medio con postcontraste y la de la derecha con el mapa de segmentación.</i>	20
<i>Figura 2. Conjunto de imágenes.</i>	20
<i>Figura 3. Ejemplo de arquitectura UNet.</i>	22
<i>Figura 4. Ejemplo de arquitectura GAN.</i>	23
<i>Figura 5. Arquitectura Modelo Bi-U-Net.</i>	26
<i>Figura 6. Arquitectura Modelo Bi-GAN.</i>	27
<i>Figura 7. Métricas basadas en píxeles.</i>	31
<i>Figura 8. Comparación visual entre el modelo base (U-Net) y los modelos propuestos en este trabajo (Bi-U-Net).</i>	33
<i>Figura 9. Comparación visual entre el modelo base (GAN) y los modelos propuestos en este trabajo (Bi-GAN).</i>	34
<i>Figura 10. Loss Modelos U-Net.</i>	36
<i>Figura 11. Loss Modelos GANs.</i>	36

Lista de tablas

	Pág.
Tabla 1. <i>Distribución de los datos en función de subconjunto y de tipo de lesión.</i>	21
Tabla 2. <i>Métricas de evaluación obtenidas por todos los modelos. Valores en negrilla significan los mejores resultados por cada tipo de modelo.</i>	32

Resumen

DESARROLLO DE UN MODELO COMPUTACIONAL PARA LA GENERACIÓN DE IMÁGENES ARTIFICIALES POST-CONTRASTE DE MRI ORIENTADO A LA SÍNTESIS DE REGIONES CON POSIBLES LESIONES TUMORALES EN TEJIDOS DE MAMA

JULIANA MONTOYA DUQUE

El cáncer de mama, una de las principales causas de mortalidad en mujeres a nivel mundial, presenta retos significativos en cuanto a su detección temprana y tratamiento. Este proyecto aborda dichos desafíos mediante el desarrollo de un modelo computacional avanzado para la generación de imágenes artificiales post-contraste de resonancia magnética (MRI), orientado específicamente a la síntesis de regiones con posibles lesiones tumorales en tejidos mamarios. La propuesta no solo busca mejorar la calidad del diagnóstico, sino también reducir la dependencia de agentes de contraste y mitigar riesgos asociados a su uso prolongado. La metodología del trabajo se basa en dos enfoques principales: modelos U-Net y redes generativas adversarias (GANs). Para todos los modelos, se introdujo un encoder adicional para combinar la segmentación de lesiones con información global de las imágenes. Además se implementaron 3 estrategias avanzadas de conexión entre espacios latentes para estabilizar el entrenamiento y mejorar la calidad de las imágenes generadas. Las estrategias exploradas incluyen conexiones a través de mapas de características profundas (configuración F), combinaciones lineales de espacios latentes (configuración M) y combinaciones no lineales a través de capas convolucionales (configuración B). Todas estas, orientadas a preservar la estructura y el detalle anatómico de las imágenes y las características de las lesiones tumorales. Los resultados demuestran que las variaciones propuestas (F, M y B) en ambos tipos de modelo superan a las configuraciones U-Net y GAN baseline. Las evaluaciones cuantitativas con métricas como MAE, SSIM y PSNR reflejan una mayor capacidad de las variaciones para preservar características críticas en las imágenes generadas, tanto a nivel global como en regiones de interés (ROI) con tumores. La variación B destacó consistentemente, logrando la mejor optimización y estabilidad durante el entrenamiento, especialmente en las GANs. Estos resultados demuestran que los modelos propuestos permiten generar imágenes sintéticas realistas

y de mayor fidelidad en función de tumores en tejidos mamarios. La integración de este trabajo en sistemas de apoyo al diagnóstico radiológico, podría fortalecer los sistemas en optimización de recursos y reducción de riesgo en los pacientes.

Palabras claves: Cáncer de mama, síntesis de imágenes, deep learning, imágenes de resonancia Magnética contrastada (DCE-MRI)

Abstract

JULIANA MONTOYA DUQUE

Breast cancer, one of the main causes of mortality in women worldwide, presents significant challenges in terms of its early detection and treatment. This project addresses these challenges by developing an advanced computational model for the generation of artificial post-contrast magnetic resonance imaging (MRI) images, specifically aimed at the synthesis of regions with possible tumor lesions in breast tissues. The proposal not only seeks to improve the quality of diagnosis, but also reduce dependence on contrast agents and mitigate the risks associated with their prolonged use. The work methodology is based on two main approaches: U-Net models and generative adversarial networks (GANs). For all models, an additional encoder was introduced to combine lesion segmentation with global image information. In addition, 3 advanced connection strategies between latent spaces were implemented to stabilize the training and improve the quality of the generated images. Strategies explored included connections via deep feature maps (configuration F), linear combinations of latent spaces (configuration M), and non-linear combinations via convolutional layers (configuration B). All of these, aimed at preserving the structure and anatomical detail of the images and the characteristics of the tumor lesions. The results show that the proposed variations (F, M and B) in both types of model outperform the U-Net and GAN baseline configurations. Quantitative evaluations with metrics such as MAE, SSIM and PSNR reflect a greater ability of the variations to preserve critical features in the generated images, both globally and in regions of interest (ROI) with tumors. Variation B consistently stood out, achieving the best optimization and stability during training, especially in GANs. These results demonstrate that the proposed models allow the generation of realistic and higher fidelity synthetic images based on tumors in breast tissues. The integration of this work in support systems for radiological diagnosis could strengthen the systems in optimization of resources and reduction of risk in patients.

Keywords: Breast cancer, image synthesis, deep learning, dynamic contrast enhanced magnetic resonance imaging (DCE-MRI)

Glosario

Aprendizaje Computacional: Estudia la capacidad de construir sistemas que aprendan a partir de la introducción de datos[1].

Cáncer de mama: Es una enfermedad donde las células comienzan a multiplicarse sin control y acumularse en las mamas formando lesiones o tumores.

Deep Learning: Son algoritmos inspirados en las redes neuronales del cerebro humano, con ellos se busca aprender de una gran cantidad de datos, hace parte de la inteligencia artificial.

Resonancia magnética: Implica el uso de un potente imán y de ondas de radio que permiten observar las estructuras del cuerpo humano y sus órganos.

Introducción

El cáncer de mama se ha convertido en un tema de gran relevancia en el ámbito de la salud pública a nivel mundial, siendo la principal causa de mortalidad entre las mujeres. Este tipo de cáncer, que se origina por la multiplicación descontrolada de células en el tejido mamario, es el más común entre las mujeres y, lamentablemente, también es responsable de una alta tasa de fallecimientos, con aproximadamente 670,000 muertes registradas en 2022. La detección temprana de esta enfermedad es fundamental, ya que las posibilidades de supervivencia son considerablemente mayores cuando se diagnostica en etapas iniciales.

La resonancia magnética con contraste (DCE-MRI) se ha consolidado como una de las herramientas más efectivas para la detección del cáncer de mama, gracias a su alta sensibilidad diagnóstica. Sin embargo, esta técnica presenta desafíos significativos, como la dependencia de la interpretación por parte de radiólogos y el uso de medios de contraste que pueden ser perjudiciales para la salud. Esto ha llevado a la búsqueda de alternativas que no solo mejoran la precisión del diagnóstico, sino que también minimicen los riesgos asociados.

En este contexto, la inteligencia artificial ha comenzado a jugar un papel crucial. La posibilidad de desarrollar modelos computacionales que generen imágenes post-contraste de resonancia magnética podría transformar la forma en que se diagnostica el cáncer de mama. Estos modelos no solo permitirían una identificación más precisa de lesiones tumorales, sino que también reducirían la carga sobre los especialistas y los costos asociados al diagnóstico. Este proyecto se centra en la creación de un modelo que sintetice imágenes de alta calidad, contribuyendo así a la detección temprana del cáncer de mama y facilitando el acceso a herramientas diagnósticas más seguras y efectivas para las pacientes.

1. Planteamiento del problema

1.1 Descripción

El cáncer de mama es un problema de salud pública a nivel mundial, además de ser la principal causa de mortalidad en mujeres. Es el primer tipo de cáncer con mayor tasa de incidencias y el quinto con mayor tasa de mortalidad en el mundo [2]. A pesar de que la tasa de supervivencia depende de la etapa de la enfermedad al momento de su diagnóstico, pacientes con la enfermedad en etapa temprana poseen una tasa de supervivencia a 5 años significativamente más elevada que aquellos quienes han sido diagnosticados en etapa avanzada (91% en etapa temprana versus 29% en etapa avanzada) [3]

En la actualidad, existen diversos tipos de exámenes médicos que se utilizan para el diagnóstico de lesiones y de cáncer. Entre ellas, la técnica con mayor sensibilidad diagnóstica para la detección de la enfermedad es el de las imágenes potenciadas por medios de contraste en resonancia magnética (DCE-MRI por sus siglas en inglés) [4]. En ella, usualmente se toma la proyección del tejido previo a la aplicación del medio de contraste (imagen pre-contraste) y se toman múltiples proyecciones en diferentes instantes de tiempo después de la aplicación del medio de contraste (imágenes post-contraste) A pesar de ello, sus mayores limitaciones son que está estrictamente ligada a la interpretación de radiólogos y personal especializado en este tipo de imagen diagnóstica, situación que aumenta los costes por personal y limita la cantidad de estudios analizados por unidad de tiempo, además de que los medios de contraste son componentes químicos tóxicos para la salud humana.

Recientemente, la implementación de estrategias basadas en inteligencia artificial para el apoyo diagnóstico del cáncer de mama ha tomado más fuerza debido al potencial en procesamiento computacional disponible en la actualidad, que a su vez permite el análisis de imágenes y estudios completos en tiempo real, sin depender exclusivamente de la subjetividad de los especialistas reduciendo los costos operativos del proceso [5]. Además, también ha mostrado resultados promisorios en la generación de imágenes con valor diagnóstico, con el fin de reducir los riesgos biológicos asociados al uso de medios de contraste [6]. En este sentido, este proyecto

tiene como objetivo desarrollar un modelo computacional que permita generar imágenes post-contraste de MRI orientado a la síntesis de las lesiones malignas o benignas. De esta manera, el modelo podrá generar imágenes con identificación visual de captación de contraste que permita identificar lesiones en el tejido. Se espera que este proyecto contribuya significativamente al diagnóstico temprano del cáncer de mama a largo plazo, además de contribuir con la inclusión de herramientas generativas en el ámbito médico.

1.2 Formulación

¿Podría entrenarse un modelo computacional que genera imágenes artificiales post-contraste de resonancia magnética (MRI), enfocado en la síntesis de regiones que podrían contener lesiones tumorales en tejidos mamarios?

2. Justificación

El cáncer de mama representa una de las principales causas de mortalidad en mujeres a nivel mundial, afectando tanto a países desarrollados como en desarrollo. La detección temprana de este tipo de cáncer es crucial para aumentar las tasas de supervivencia y mejorar la calidad de vida de las pacientes. En este contexto, el desarrollo de un modelo computacional avanzado que sea capaz de generar imágenes sintéticas de alta calidad, enfocadas en la identificación precisa de regiones con posibles lesiones tumorales en tejidos mamarios, podría revolucionar las herramientas actuales de diagnóstico. Estas imágenes sintéticas no solo facilitarían la identificación temprana de tumores, sino que también reducirían la exposición de las pacientes a agentes de contraste intravenosos, que a menudo contienen metales como el gadolinio. Dichos elementos, aunque útiles en la obtención de imágenes diagnósticas, pueden acumularse en ciertos órganos del cuerpo humano, como el cerebro y los riñones, incrementando el riesgo de complicaciones a largo plazo. Por tanto, una solución computacional eficiente y segura podría ser un complemento valioso en el entorno médico, proporcionando un diagnóstico más accesible y menos invasivo.

Actualmente, la disponibilidad de datasets de imágenes de resonancia magnética (MRI) post-contraste etiquetadas y de alta calidad es extremadamente limitada, lo cual representa un desafío significativo en el avance de las herramientas de diagnóstico basadas en inteligencia artificial. Estas limitaciones surgen principalmente debido a restricciones relacionadas con la privacidad de los pacientes y los altos costos asociados con la recopilación, procesamiento y almacenamiento de datos clínicos reales [7]. Al desarrollar un modelo computacional capaz de generar imágenes artificiales post-contraste con un nivel elevado de realismo y características anatómicas fieles, se podrían superar estas barreras, reduciendo significativamente la dependencia de conjuntos de datos clínicos reales. Esto no solo aumentaría la disponibilidad de datos para la investigación, sino que también abriría nuevas oportunidades para entrenar y validar algoritmos avanzados de detección y segmentación de tumores en un entorno controlado y reproducible. Además, esta tecnología facilitaría la colaboración interdisciplinaria entre investigadores y profesionales médicos, impulsando el desarrollo de herramientas diagnósticas más accesibles, eficaces y éticamente responsables.

3. Objetivos

3.1 Objetivo general

Desarrollar un modelo computacional para la generación de imágenes artificiales post-contraste de MRI orientado a la síntesis de regiones con posibles lesiones tumorales en el tejidos de mama.

3.2 Objetivos específicos

- Recopilar y preprocesar una base de datos pública que contenga imágenes de DCE-MRI con anotaciones y segmentaciones disponibles para el desarrollo de modelos computacionales.
- Diseñar y entrenar un algoritmo de síntesis de imágenes post-contraste de resonancia magnética (MRI) que reproduzca características visuales propias de regiones con hallazgos clínicos relevantes.
- Validar la precisión y la sensibilidad del modelo computacional mediante pruebas comparativas con los conjuntos de anotaciones en el tejido mamario.

4. Marco teórico

Antecedentes

En los últimos años, se han introducido nuevos métodos de diagnóstico para validar la sospecha de cáncer, destacando la tomografía axial computada (TAC) con contraste y la resonancia magnética (RM), que son más sensibles al detectar la angiogénesis relacionada con lesiones malignas. Actualmente, se considera que la RM con medio de contraste intravenoso es el método más sensible para el diagnóstico y la evaluación de la extensión del cáncer de mama. Las indicaciones para realizar una RM mamaria incluyen el estudio de la extensión en cánceres ya diagnosticados, la sospecha de recidiva tumoral, el tamizaje en pacientes con alto riesgo, la evaluación de cáncer con exámenes previos inconclusos, la búsqueda de neoplasias ocultas y el seguimiento del tratamiento neoadyuvante.[8]

La imagenología por resonancia magnética (IRM) ofrece una perspectiva in vivo sobre la fisiopatología de la esclerosis múltiple (EM), considerándola una herramienta fundamental para el diagnóstico, manejo y decisiones sobre el inicio o modificación de tratamientos con fármacos modificadores de la enfermedad (FARME). Es crucial que los profesionales estén familiarizados con las técnicas de imagen y los protocolos recomendados, ya que esto puede mejorar la detección de la actividad y facilitar un manejo óptimo. A menudo, se presenta una desconexión entre clínicos y radiólogos respecto a los requisitos al solicitar un estudio de imagen, lo que puede resultar en un manejo subóptimo. Esta situación puede deberse a una comunicación deficiente, a la falta de conocimiento sobre técnicas de imagen por parte de los clínicos, o a que los radiólogos no siempre comprenden la relevancia de las características específicas de la IRM en el proceso de toma de decisiones del neurólogo tratante. Por ello, es fundamental establecer protocolos estandarizados de resonancia magnética para que tanto médicos como radiólogos se familiaricen con ellos, lo que contribuirá a mejorar la calidad de las imágenes, la detección de lesiones y la elaboración de informes radiológicos.[9]

La resonancia magnética con contraste dinámico (DCE-MRI) ha demostrado ser altamente sensible para el diagnóstico del cáncer de mama. No obstante, existen pocos algoritmos asistidos

por computadora que se enfoquen en el uso de imágenes DCE-MRI para este propósito, lo que se debe a la escasez de conjuntos de datos DCE-MRI accesibles al público.[10]

La tomosíntesis digital de mama (DBT) es una técnica de imagen por rayos X que permite la creación de volúmenes tridimensionales de los senos. Esta modalidad se utiliza principalmente para la detección del cáncer y se prevé que reemplace a la mamografía convencional en los próximos años [11]. La única forma de reducir la morbilidad y mortalidad asociadas a las lesiones sólidas en las mamas es mediante su detección temprana y un manejo adecuado y oportuno. La resonancia magnética presenta una alta sensibilidad diagnóstica para el cáncer de mama, lo que hace necesario evaluar su efectividad en el diagnóstico de mujeres con lesiones sólidas en las mamas.[12]

Las redes neuronales convolucionales han emergido como una técnica prometedora para la identificación y segmentación automática de lesiones. Este estudio tiene como objetivo diseñar e implementar una arquitectura de red neuronal convolucional que permita detectar y segmentar eficientemente las lesiones en imágenes de resonancia magnética, ajustando los hiperparámetros o modificando el enfoque propuesto en arquitecturas existentes.[13]

A través del aprendizaje profundo, la inteligencia artificial ha comenzado a convertirse en una herramienta complementaria para los especialistas en radiología, ayudando a establecer diagnósticos y minimizando la posibilidad de errores humanos. En la última década, la IA ha experimentado un crecimiento exponencial, proporcionando un sólido apoyo a los profesionales de la salud. Al emplear algoritmos y datos, estas tecnologías son capaces de identificar patrones y ofrecer conclusiones que facilitan tareas sencillas, como el monitoreo de la salud, la gestión de registros médicos, el diseño de tratamientos e incluso consultas digitales.[14]

La disponibilidad de dataset de resonancia magnética (MRI) post-contraste etiquetados y de alta calidad es limitada debido a restricciones de privacidad y costos asociados con la recopilación de datos clínicos.[15]

Al profundizar en el impacto y el alcance del aprendizaje automático en el ámbito de la

oncología mamaria, se puede vislumbrar un futuro en el que la tecnología juegue un papel fundamental en la lucha contra esta enfermedad, ofreciendo esperanza tanto a las pacientes como a la comunidad médica. Para esta investigación, se adoptó un enfoque mixto que combina datos numéricos y estadísticas con la búsqueda de patrones y relaciones causales mediante métodos estadísticos. Desde la perspectiva cualitativa, también se recurrió a datos no numéricos, como observaciones, para captar la complejidad y la riqueza de las experiencias en países que están a la vanguardia en la implementación de estas tecnologías.[16]

Imágenes Pre-contraste

Se refieren a las imágenes obtenidas antes de la administración de un medio de contraste en procedimientos de diagnóstico por imagen, como la resonancia magnética (MRI). Estas imágenes iniciales son fundamentales para establecer una línea base que permite a los radiólogos y médicos comparar la morfología y la densidad de los tejidos antes y después de la aplicación del contraste [28]. El uso de medios de contraste, como el yodo o el gadolinio, mejora significativamente la visibilidad de estructuras anatómicas y patologías, facilitando así una mejor identificación y caracterización de los tejidos. En este contexto, las imágenes pre-contraste son esenciales para evaluar correctamente las alteraciones que pueden surgir tras la administración del contraste, permitiendo un diagnóstico más preciso y efectivo. Además, estas imágenes ayudan a identificar posibles reacciones adversas o complicaciones asociadas al uso del contraste, lo que es crucial para la seguridad del paciente durante el procedimiento diagnóstico [29].

Imágenes Post-contraste

Son aquellas obtenidas después de la administración de un medio de contraste en la resonancia magnética. Este tipo de imágenes es crucial para resaltar y diferenciar estructuras anatómicas y patologías, ya que el medio de contraste mejora la visibilidad de los tejidos al alterar temporalmente su densidad en las imágenes [30]. Por ejemplo, en estudios como la resonancia magnética dinámica con realce de contraste (DCE-MRI), el uso de agentes como el gadolinio permite una mejor caracterización de lesiones, facilitando así la detección temprana de

condiciones como el cáncer. Las imágenes post-contraste permiten a los radiólogos comparar la captación del medio en diferentes tejidos, lo que es fundamental para diagnosticar y evaluar la gravedad de diversas patologías [31]. Además, estas imágenes son esenciales para monitorear posibles reacciones adversas al medio de contraste y para asegurar la seguridad del paciente durante el procedimiento.

Algoritmo de Síntesis de Imágenes Post-Contraste

Un modelo fundamentado en el aprendizaje profundo, que teniendo en cuenta las áreas de realce de contraste, produce imágenes sintéticas de DCE-MRI postcontraste a partir de imágenes precontraste. Esto podría respaldar la reducción del uso de medios contrastados en las pruebas diagnósticas de cáncer de mama [17]. Implementamos modelos de aprendizaje profundo en el proyecto para realizar la síntesis de imágenes post-contraste de MRI para la detección de las lesiones de mama.

Autoencodes

Son un tipo de red neuronal artificial que se utiliza principalmente en el contexto del aprendizaje no supervisado, con el objetivo de aprender representaciones eficientes de datos. Su estructura se compone de dos partes fundamentales: el codificador, que transforma los datos de entrada en una representación comprimida, y el decodificador, que reconstruye los datos originales a partir de esta representación. Este proceso permite a los autoencoders capturar las características más relevantes de los datos, facilitando tareas como la reducción de dimensionalidad, la eliminación de ruido y la detección de anomalías. Durante su entrenamiento, el modelo ajusta sus parámetros para minimizar la diferencia entre los datos de entrada y su reconstrucción, optimizando así su capacidad para identificar patrones subyacentes. Los autoencoders han encontrado aplicaciones en diversas áreas, incluyendo la compresión de imágenes, la generación de datos sintéticos y la mejora de sistemas de recomendación, lo que los convierte en herramientas versátiles y potentes en el campo del aprendizaje automático y la inteligencia artificial.

5. Metodología

5.1 Tipo de proyecto

La estrategia metodológica implementada en el proyecto será investigación.

5.2 Método

El estudio se llevó a cabo utilizando imágenes con tumores extraídas de la base de datos de estudios prospectivos de resonancia magnética presentado en Zhao et al. [21]. Este conjunto de datos consta de 232 pacientes seleccionados con casos benignos y malignos. Cada caso consta de tres tipos de secuencias: secuencias precontraste, poscontraste y sustracción. Además, la implementación de los algoritmos de procesamiento y aprendizaje se realizó en los servidores de la Institución Universitaria Pascual Bravo, los cuales cuentan con 16 CPUs procesamiento, 32 GB en memoria RAM y tarjeta gráfica RTX A4000, empleando Visual Studio Code como entorno de desarrollo. A continuación, se presentan las actividades realizadas y los productos esperados para alcanzar los objetivos planteados.

1. Recopilación de datos

Para la recopilación de datos se trabaja con una base de datos ligada a la investigación anteriormente mencionada. Esta base de datos consta con casos benignos y malignos, la base de datos se divide en imágenes pre-contraste, post-contraste y sustracción, y el diagnóstico realizado a los pacientes. Inicialmente, esta información es aprovechada para crear un archivo de metadata, con el fin de construir un dataloader eficiente para validación de información. Con ello se crean 3 subconjuntos: entrenamiento (train), prueba (test) y validación (val) consolidados en 3 archivos .csv con las rutas relativas donde se encuentran las imágenes dentro de la base de datos, esto hará que a la hora de cargarlo en el dataloader la carga de las imágenes sea más eficiente y menor el gasto computacional. El dataloader consta de un Input en este caso es la imagen pre-contraste y dos outputs correspondientes a la imagen post-contraste y el mapa de segmentación del tumor. La figura 1 muestra un ejemplo de las imágenes mencionadas.

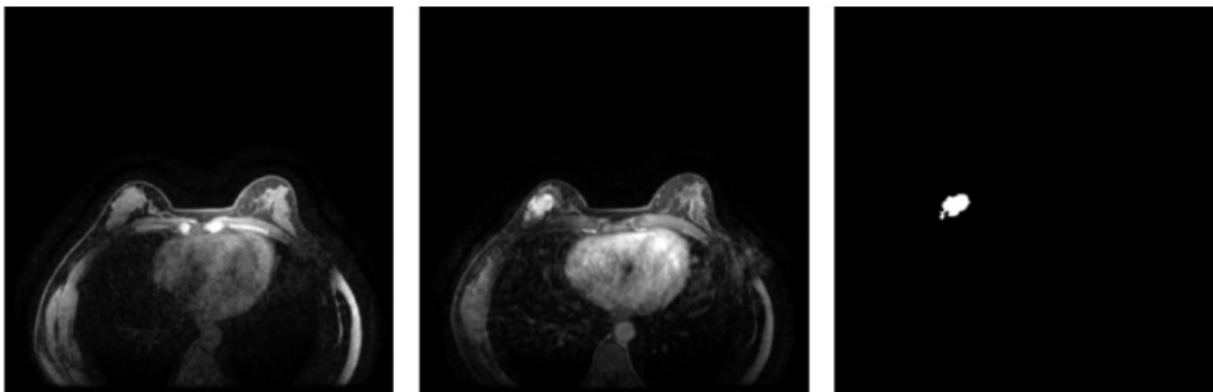


Figura 1: *Imágenes de ejemplo extraídas de la base de datos. La imagen de la izquierda corresponde con la precontraste, la de en medio con postcontraste y la de la derecha con el mapa de segmentación. Fuente : Elaboración propia*

Las imágenes se reescalaron a un tamaño de 256x256 píxeles para facilitar su procesamiento mediante interpolación bicúbica, garantizando así la preservación de su calidad visual. Luego, fueron normalizadas al rango de -1 a 1 para optimizar su compatibilidad con los modelos de procesamiento. Adicionalmente, se transformaron a escala de grises, lo que permitió reducir la complejidad computacional del análisis. El conjunto de datos procesado se empleó para el entrenamiento de los modelos. En la Figura 2 se ilustra un ejemplo de la distribución de las imágenes obtenidas. Además, la tabla 1 especifica la cantidad de imágenes por subconjunto y tipo de lesión

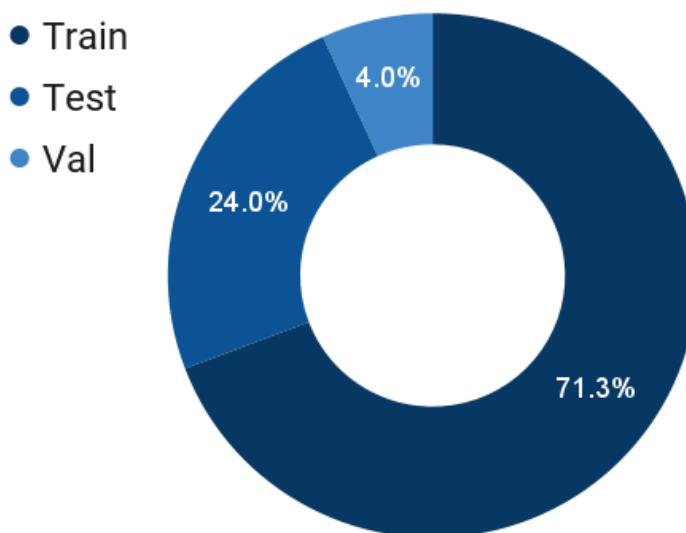


Figura 2: *Conjunto de imágenes. Fuente : Elaboración propia*

Tabla 1: *Distribución de los datos en función de subconjunto y de tipo de lesión.*

Imágenes	Lesiones Benignas	Lesiones Malignas
Train	327	875
Test	128	289
Validation	24	43
Total Imágenes	479	1207

2. Modelos de Aprendizaje

2.1 Modelos Base

Para esta etapa se emplean algoritmos de aprendizaje profundo para síntesis de imágenes como los autoencoders en configuración U-Net y modelos adversarios generativos (GAN).

Autoencoder tipo U-Net

El modelo U-Net es una arquitectura de red neuronal convolucional diseñada para tareas de segmentación de imágenes, particularmente útil en aplicaciones como imágenes médicas. Propuesto en 2015 por Olaf Ronneberger y su equipo, su estructura se basa en un codificador (encoder) que extrae características mediante capas convolucionales y operaciones de reducción de resolución (*downsampling*), y un decodificador (decoder) que reconstruye la segmentación incrementando la resolución (*upsampling*). Las conexiones en U (*skip connections*) combinan características del codificador y decodificador en cada nivel, preservando detalles espaciales esenciales como bordes. Entre sus principales ventajas destacan su eficiencia en conjuntos de datos pequeños y su capacidad para detectar bordes de manera precisa, aprovechando tanto el contexto global como los detalles locales.

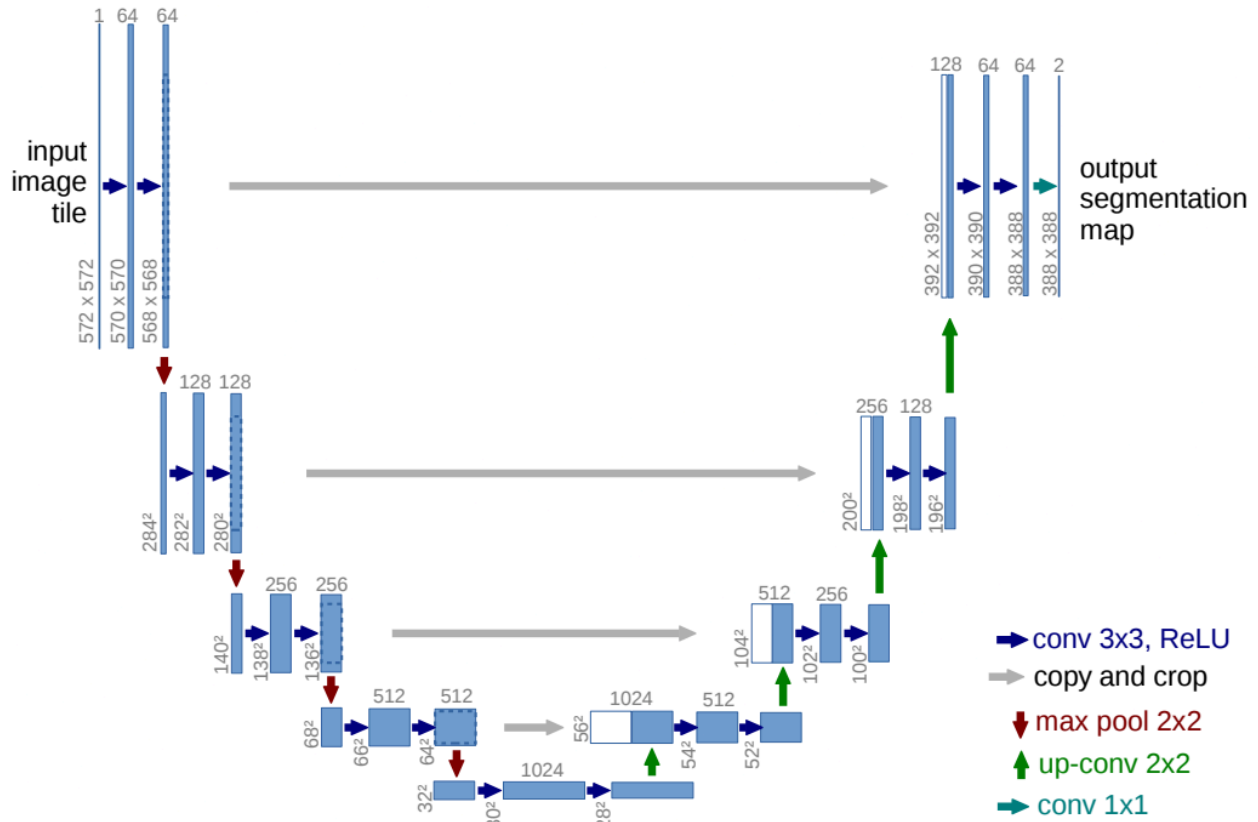


Figura 3: Ejemplo de arquitectura UNet. Fuente: Ronneberger et al. [22]

Error Absoluto Medio (Mean Absolute Error, MAE) como función de costo

Para optimizar los modelos U-Net se emplea el Error Absoluto Medio (Mean Absolute Error, MAE) como métrica de optimización en los píxeles. El MAE es una métrica que mide el promedio de las diferencias absolutas entre los valores predichos y los valores reales. Es ampliamente utilizada como función de costo en redes neuronales, incluidas arquitecturas como U-Net, debido a su simplicidad y capacidad para evaluar errores de manera directa. En el contexto de una U-Net, el MAE se define como:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

- **Cross-Entropy Loss:** Evalúa la diferencia entre las predicciones del modelo y las etiquetas verdaderas, considerando cada píxel como una probabilidad. Es adecuada para segmentación binaria o multiclase.

$$\text{Loss} = -\frac{1}{N} \sum_{i=1}^N y_i \log(\hat{y}_i)$$

Generative Adversarial Networks (GANs)

El modelo Generative Adversarial Networks (GANs) es una arquitectura de red neuronal diseñada para generar datos nuevos y realistas a partir de ruido aleatorio, ampliamente utilizada en tareas como generación de imágenes, videos, y síntesis de datos. Propuesto en 2014 por Goodfellow et al, [23]. Su estructura se basa en dos redes neuronales que compiten entre sí: un generador que crea datos sintéticos y un discriminador que evalúa si los datos son reales o generados. Durante el entrenamiento, el generador mejora creando datos más convincentes, mientras que el discriminador afina su capacidad para distinguirlos, en un proceso adversarial. Entre las principales ventajas de las GANs destacan su capacidad para aprender distribuciones complejas de datos y su aplicabilidad en problemas donde los datos etiquetados son limitados, aunque su entrenamiento puede ser inestable y requiere un balance cuidadoso entre las dos redes. La figura 4 muestra un ejemplo de arquitectura GAN.

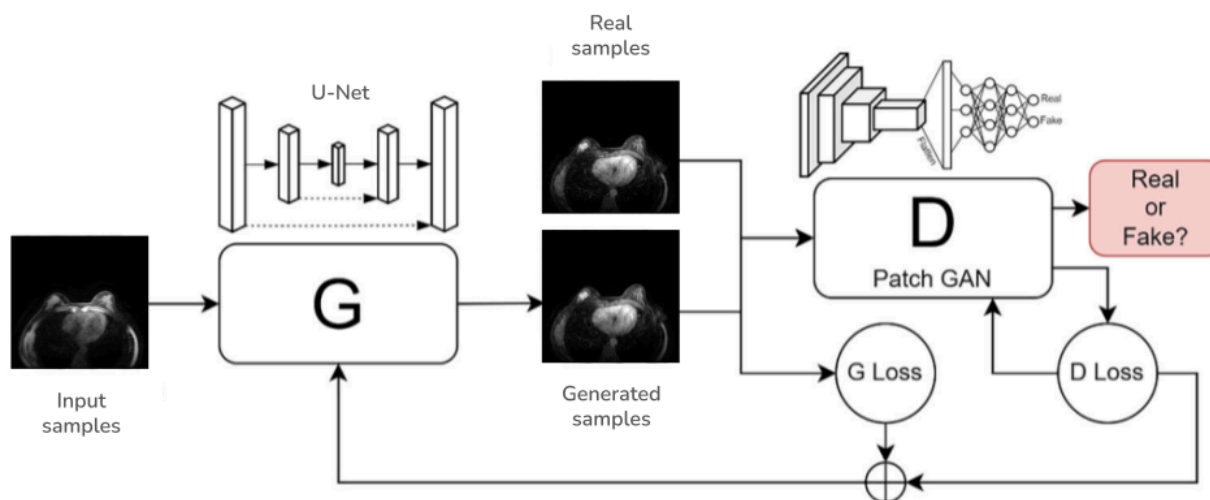


Figura 4: Ejemplo de arquitectura GAN. Fuente: Adiyaman et al. [24]

Entropía cruzada y función adversaria como Funciones de Costo en GANs

La entropía cruzada es una función de costo ampliamente utilizada en redes neuronales, especialmente en modelos GANs. Mide la disimilitud entre las distribuciones reales y predichas, penalizando las predicciones incorrectas con mayor severidad. En GANs, esta

función se emplea tanto para el generador como para el discriminador, ayudando a evaluar cómo de bien el generador puede engañar al discriminador y cómo de efectivo es el discriminador al distinguir datos reales de sintéticos. Para un problema binario, la función de costo de entropía cruzada se define como:

$$H(y, \hat{y}) = -\frac{1}{N} \sum_{i=1}^N (y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i))$$

Donde N es el número total de muestras, y_i es la etiqueta real de la i -ésima muestra (1 para datos reales, 0 para datos sintéticos) y \hat{y}_i es la probabilidad predicha de que la i -ésima muestra sea real.

En este modelo, el discriminador (D) se entrena para diferenciar entre **Datos reales** (provenientes de la distribución de datos reales) y **Datos generados** (producidos por el generador a partir de una distribución original). Su objetivo es maximizar la probabilidad de asignar correctamente datos reales ($y=1$) y datos generados ($y=0$). La función de costo para el discriminador se define como:

$$\mathcal{L}_D = -\mathbb{E}_{x \sim p_{\text{real}}} [\log(D(x))] - \mathbb{E}_{z \sim p_z} [\log(1 - D(G(z)))]$$

Por otro lado, El generador (G) busca engañar al discriminador para que clasifique sus muestras generadas como reales. Su objetivo es maximizar las predicciones del discriminador que clasifiquen los datos generados como reales ($y=1$), que corresponde al caso donde el discriminador falla en distinguir entre datos reales y generados.

$$\mathcal{L}_G = -\mathbb{E}_{z \sim p_z} [\log(D(G(z)))]$$

La interacción entre G y D puede expresarse como un problema de minimización para el generador y maximización para el discriminador, formulado como:

$$\min_G \max_D \mathbb{E}_{x \sim P_{\text{real}}} [\log(D(x))] + \mathbb{E}_{z \sim P_z} [\log(1 - D(G(z)))]$$

En este contexto, D intenta maximizar la función para clasificar correctamente los datos reales y falsos; y G intenta minimizar la función para generar muestras indistinguibles de los datos reales. Este entrenamiento adversarial crea un equilibrio donde G aprende a generar datos más realistas, aproximando la distribución de datos reales, mientras que D mejora su capacidad de discriminación.

Cuando G y D están perfectamente entrenados, $D(G(z)) \approx 0.5$ para datos generados, indicando que el discriminador ya no puede distinguir entre datos reales y falsos. En este punto, G ha alcanzado su objetivo. Sin embargo, esta dinámica es lo que hace a las GANs poderosas pero también difíciles de entrenar, requiriendo ajustes cuidadosos en arquitecturas y tasas de aprendizaje.

2.2 Modelos Propuestos

Arquitectura Bi-U-Net

Para condicionar el modelo utilizando la información de los tumores presentes en la base de datos, se diseñó una arquitectura basada en U-Net que incluye, además del encoder convencional (E) un encoder adicional denominado que procesa los mapas de tumor (R). Este encoder recibe como entrada la imagen de la lesión y la procesa para codificar en un vector de representación latente. El propósito de este vector es capturar las características globales de las imágenes, además de las características distintivas de la lesión, permitiendo integrarlas con los componentes extraídos del proceso de segmentación de la imagen completa. La combinación de ambos componentes se realiza para enriquecer la representación de los datos y asegurar que la información relevante de la lesión esté correctamente incorporada. Una vez realizada esta integración, un decodificador (D) toma como entrada los datos mezclados y los utiliza para reconstruir la imagen completa, preservando tanto la información contextual como los detalles específicos de la lesión segmentada. Esta arquitectura denominada **Bi-U-Net** permite aprovechar la capacidad de la U-Net para realizar segmentaciones precisas, mientras que el encoder adicional asegura que las características del tumor sean explícitamente consideradas durante el proceso. En la Figura 5 se presenta una ilustración detallada de la arquitectura propuesta, destacando cómo el flujo de información es procesado y combinado en cada etapa del modelo. Este diseño innovador busca optimizar la segmentación y reconstrucción de imágenes médicas en escenarios donde los tumores son el enfoque principal. Adicionalmente, La función de costo de la Bi-U-Net se emplea el MAE en los píxeles, tal y como en la implementación original. Sin embargo, los parámetros de los encoders E y R se ajustan de manera independiente.

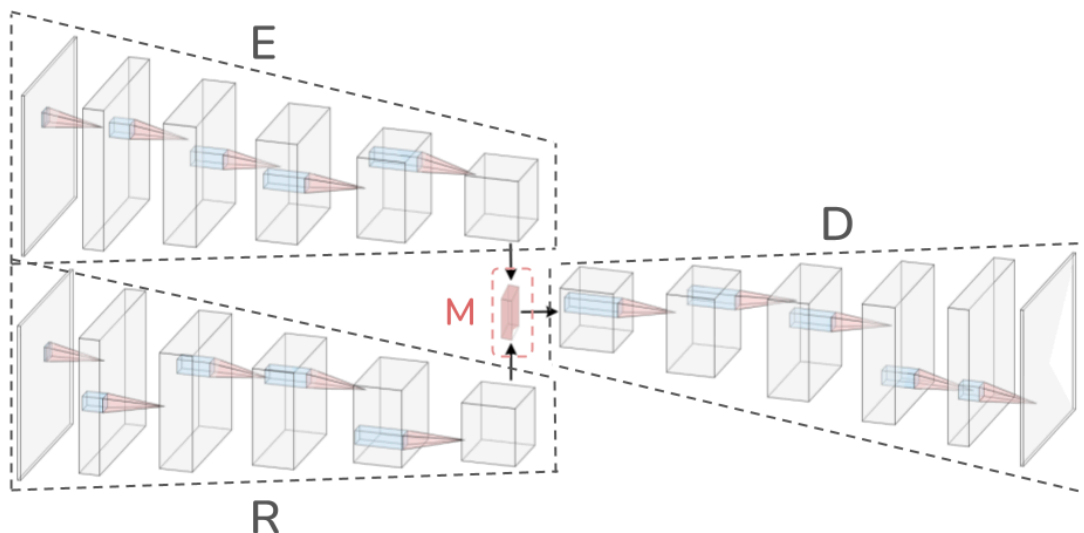


Figura 5: *Arquitectura Modelo Bi-U-Net. Fuente: Elaboración propia.*

Arquitectura Bi-GAN

Con el fin de evaluar el desempeño en cuanto a realismo se emplea la misma arquitectura planteada anteriormente pero además se utiliza la función adversaria para entrenar el modelo.

Con el objetivo de mejorar la síntesis de imágenes condicionadas por los tumores, se empleó una arquitectura basada en la Bi-U-Net en configuración de GAN (Bi-GAN). Sin embargo, se añade un discriminador, por su parte, está diseñado para evaluar la calidad de las imágenes reconstruidas por el generador. Su tarea principal es diferenciar entre las imágenes reconstruidas y las imágenes reales de la base de datos, forzando al generador a producir salidas más realistas y consistentes. Este proceso de retroalimentación adversarial fomenta una mejora progresiva del generador, ayudando a capturar tanto los detalles finos de las lesiones como las características globales de la imagen.

En la Figura 6 se muestra un esquema detallado de la arquitectura propuesta, destacando las interacciones clave entre el generador y el discriminador, así como el flujo de datos en cada componente. Este diseño representa un enfoque avanzado para la segmentación y reconstrucción de imágenes médicas, aprovechando las capacidades complementarias de las U-Net y las GANs. En este caso, la función de costo para este modelo se compone de la función adversaria junto con el MAE como función de pérdida de los píxeles. Los encoders E y R se entrenan de manera independiente, al igual que en la Bi-U-Net.

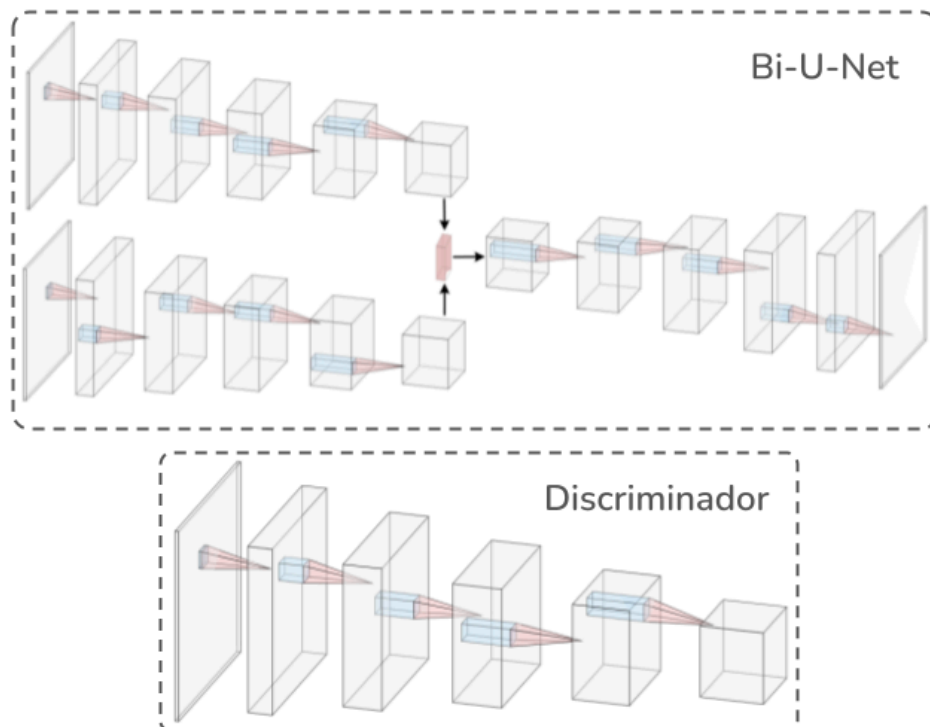


Figura 6: *Arquitectura Modelo Bi-GAN. Fuente: Elaboración propia.*

Combinación de espacios latentes

Con el objetivo de identificar la estrategia más efectiva para combinar los espacios latentes generados por los encoders E y R, se han propuesto múltiples enfoques que buscan maximizar la integración de las características extraídas. Estas estrategias incluyen desde combinaciones simples, como la concatenación directa de los vectores latentes, hasta técnicas más avanzadas que emplean operaciones elementales (como suma o multiplicación) o arquitecturas específicas para fusionar información, con el fin de preservar tanto las relaciones estructurales globales como los detalles locales capturados por cada encoder.

Conexión a través de mapas de características profundas (F): En esta implementación, se combinan los mapas de características extraídos por ambos encoders y decoders mediante una operación de concatenación, implementada a través de *skip connections* convencionales. Estas conexiones permiten una integración eficiente de las características globales de ambas imágenes con las características locales capturadas por los mapas de segmentación de los tumores. Este enfoque no solo facilita la preservación de información clave en diferentes niveles de resolución, sino que también mejora la transferencia de detalles estructurales y contextuales, optimizando el proceso de segmentación y reconstrucción final.

Conexión a través multiplicación de espacios latentes (M): En este enfoque, los espacios latentes generados por los encoders son combinados mediante una operación de multiplicación elemento a elemento entre ambos componentes. Este procedimiento busca lograr una combinación lineal de los espacios latentes, permitiendo resaltar características comunes o

interacciones clave entre las representaciones latentes obtenidas de las imágenes procesadas. A diferencia de la concatenación, la multiplicación introduce una forma más selectiva de fusión, donde solo las características que coinciden o se refuerzan entre los espacios contribuyen significativamente a la representación final, potenciando la integración de información relevante para las tareas de segmentación y reconstrucción.

Conexión a través de combinación convolucional (B): En este enfoque, los bloques de características obtenidos de los espacios latentes de ambos encoders son primero concatenados y luego procesados a través de una capa convolucional. Este método permite establecer relaciones espaciales no lineales entre las componentes de los espacios latentes, capturando interacciones complejas y patrones que no serían posibles con operaciones lineales simples. La capa convolucional actúa como un mecanismo de integración, aprendiendo automáticamente cómo fusionar las características globales y locales de manera óptima, mejorando así la capacidad del modelo para representar información relevante tanto para la segmentación como para la reconstrucción.

3. Evaluación de Desempeño

Por último se evalúa el desempeño de cada algoritmo comparando funciones de costo y métricas basadas en píxeles como MAE, SSIM, PSNR.

- El **Error Absoluto Medio (MAE)** es una métrica ampliamente utilizada para cuantificar la discrepancia entre valores predichos y valores reales en diversos contextos, como la comparación de imágenes. Su objetivo es medir la diferencia promedio en términos absolutos, lo que proporciona una interpretación directa e intuitiva del error. En el contexto de imágenes, el MAE evalúa la fidelidad de una imagen reconstruida en relación con su contraparte original, calculando el promedio de las diferencias absolutas a nivel de píxel. Esto implica que cada píxel en la imagen predicha (\hat{y}) es comparado con el píxel correspondiente en la imagen real (y), sin considerar la dirección del error (positivo o negativo). El resultado es un valor escalar que representa el error promedio, donde valores más bajos indican una mayor similitud entre las dos imágenes.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

- **SSIM:** El Índice de Similitud Estructural (SSIM) es una métrica avanzada utilizada para evaluar la similitud entre dos imágenes, con un enfoque especial en las características

estructurales de la imagen. A diferencia de otras métricas como el MAE o el MSE, que solo miden las diferencias en términos de intensidad de píxel, el SSIM tiene en cuenta factores perceptuales importantes como el contraste, la luminancia y la estructura local, lo que lo convierte en una métrica más alineada con la percepción humana de la calidad visual. En lugar de centrarse solo en las diferencias absolutas de píxeles, el SSIM compara las imágenes en tres componentes fundamentales que influyen en cómo los humanos perciben las imágenes visualmente:

Luminancia (L): Mide la similitud en términos de brillo o intensidad de la imagen. Se enfoca en cómo se alinean los niveles de brillo entre las imágenes comparadas, asegurando que los detalles oscuros o brillantes sean interpretados correctamente.

Contraste (C): Evalúa las diferencias en la variabilidad de la intensidad entre las imágenes. El contraste mide cuán bien se distinguen las áreas más claras y más oscuras de las imágenes, lo que es esencial para capturar detalles importantes como bordes o texturas.

Estructura (S): Se refiere a la similitud en los patrones de estructura local, como las texturas o las formas. Este componente evalúa cómo las estructuras de alto nivel (por ejemplo, las formas de los objetos en la imagen) se alinean entre las dos imágenes.

El SSIM se formula de la siguiente manera:

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

- **PSNR (Relación Señal a Ruido de Pico):** La Relación Señal a Ruido de Pico (PSNR) es una métrica utilizada comúnmente para evaluar la calidad de una imagen reconstruida o comprimida en comparación con la imagen original. Esta métrica se basa en la relación entre la máxima potencia de la señal (la información original de la imagen) y la potencia del ruido introducido durante el proceso de reconstrucción, compresión o procesamiento. El PSNR proporciona una medida cuantitativa de la fidelidad visual de la imagen reconstruida, siendo particularmente útil en tareas de compresión de imágenes, eliminación de ruido y

restauración. El PSNR se calcula en términos de la diferencia de píxel entre las imágenes comparadas, y su fórmula se expresa como:

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{MAX}^2}{\text{MSE}} \right)$$

5.3 Instrumentos de recolección de información

5.3.1 Fuentes primarias. Artículos indexados de revistas científicas

5.3.2 Fuentes secundarias.

- Imágenes retrospectivas de la base de datos BreastDM [21].
- Material de clase, visto en la electiva - Reconocimiento de patrones y Línea de profundización - Machine Learning.

6. Resultados del Proyecto

6.1. Evaluación comparativa entre modelos

En primer lugar, se realizó una evaluación exhaustiva del desempeño de cada algoritmo, utilizando funciones de costo y métricas específicas basadas en píxeles, como el MAE (Mean Absolute Error), el SSIM (Structural Similarity Index) y el PSNR (Peak Signal-to-Noise Ratio). Estas métricas fueron seleccionadas debido a su capacidad complementaria para analizar diferentes aspectos de las imágenes generadas. Por un lado, el MAE ofrece una medida directa de la precisión numérica al cuantificar la discrepancia promedio entre los valores de píxeles en las imágenes reales y reconstruidas. Por otro lado, el SSIM y el PSNR evalúan la calidad perceptual, centrándose en la preservación de la estructura, el contraste y la luminancia, factores cruciales para garantizar la coherencia visual.

En la **Figura 7** se muestran los resultados comparativos entre las métricas SSIM y MAE, resaltando las diferencias en la capacidad de los modelos para mantener estructuras críticas y detalles relevantes en las imágenes reconstruidas. Este análisis permitió identificar patrones específicos en el desempeño de los algoritmos, destacando aquellos que priorizan la precisión estructural frente a los que maximizan la fidelidad espectral. En este plot es posible evidenciar una mejora en todas las variaciones F, M y B para ambos modelos U-Net y GAN, y posicionando al modelo U-Net como el mejor respecto a ambas métricas. De manera detallada, estas métricas son mostradas en la tabla 2, en donde además; la variación B obtiene las mejores métricas para ambos modelos.

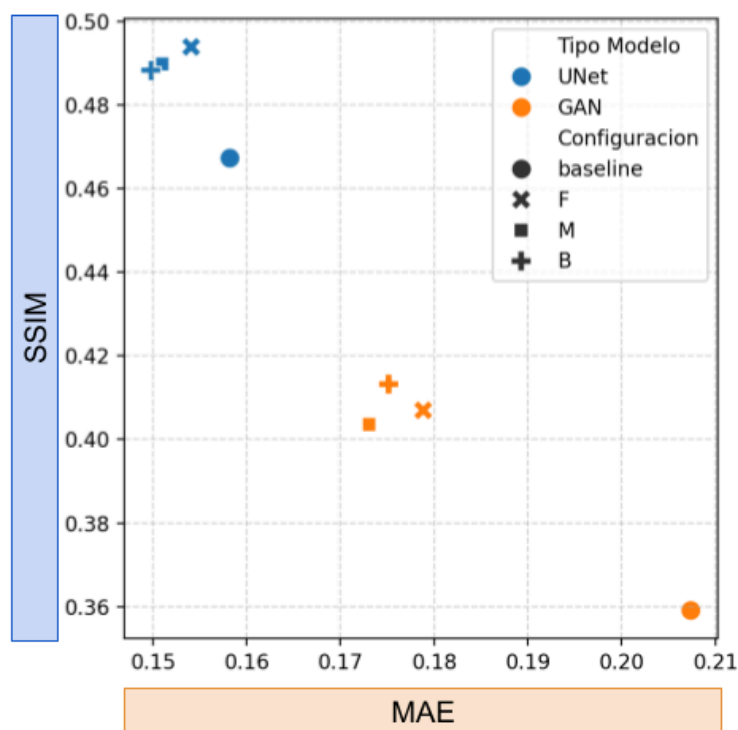


Figura 7: Métricas basadas en píxeles. Fuente: Elaboración propia.

Tabla 2: Métricas de evaluación obtenida por todos los modelos. Valores en negrilla significan los mejores resultados por cada tipo de modelo.

Tipo Modelo	Configuración	MAE	PSNR	SSIM
U-Net	baseline	0.158223	19.550144	0.467339
	F	0.154081	20.20601	0.48837
Bi-U-Net	M	0.150988	19.968097	0.489905
	B	0.149797	20.691652	0.493976
GAN	baseline	0.207405	19.394474	0.359044
	F	0.178844	20.258171	0.406976
Bi-GAN	M	0.173079	19.968965	0.403598
	B	0.175151	20.41182	0.413222

Estimación visual por los modelos Bi-U-Net

La figura muestra una comparación de resultados entre el modelo U-Net (baseline) y tres variantes del modelo Bi-U-Net según las descripciones realizadas en la sección anterior: Bi-U-Net (F), Bi-U-Net (M) y Bi-U-Net (B) en la tarea de generación de imágenes post-contraste. La primera fila corresponde con la imagen completa, donde se observa la imagen de entrada (Input), la imagen objetivo (Output) y las imágenes generadas por los diferentes modelos. La segunda fila muestra el mapa de diferencias para la imagen completa, en donde se resaltan las diferencias entre la imagen salida real y las imágenes generadas por cada modelo. La tercera fila muestra la Región de interés (ROI) que contiene una lesión tumoral por cada imagen, y por último, la cuarta fila muestra el mapa de diferencias para los ROI, resaltando las discrepancias específicas dentro del ROI tumoral entre las imágenes reales y las generadas.

Las figuras muestran múltiples hallazgos importantes. Primero, el modelo U-Net (baseline) muestra diferencias más pronunciadas (zonas intensamente rojas) en comparación con las variantes de Bi-U-Net, especialmente en regiones donde la estructura del tejido es más compleja. Además, este modelo genera una reconstrucción que, aunque capta la forma general, parece suavizada en comparación con la salida real. Esto se ve reflejado en los mapas de diferencia y el ROI de dicha columna. Sin embargo, al observar las variantes de Bi-U-Net, las diferencias poseen menor intensidad, siendo el modelo Bi-U-Net (B) el que logra una mejor aproximación con diferencias menores y distribuidas de manera homogénea en la imagen global, y los modelos Bi-U-Net (M) y Bi-U-Net (B) los que presentan una mayor contraste, fidelidad en la estructura y preservación de los detalles dentro de la región tumoral.

El modelo **Bi-U-Net (B)** demuestra ser la configuración más robusta, mostrando la menor variación en los mapas de diferencia, lo que lo posiciona como el más adecuado para

aplicaciones que requieren alta precisión en la reconstrucción de imágenes de resonancia magnética.

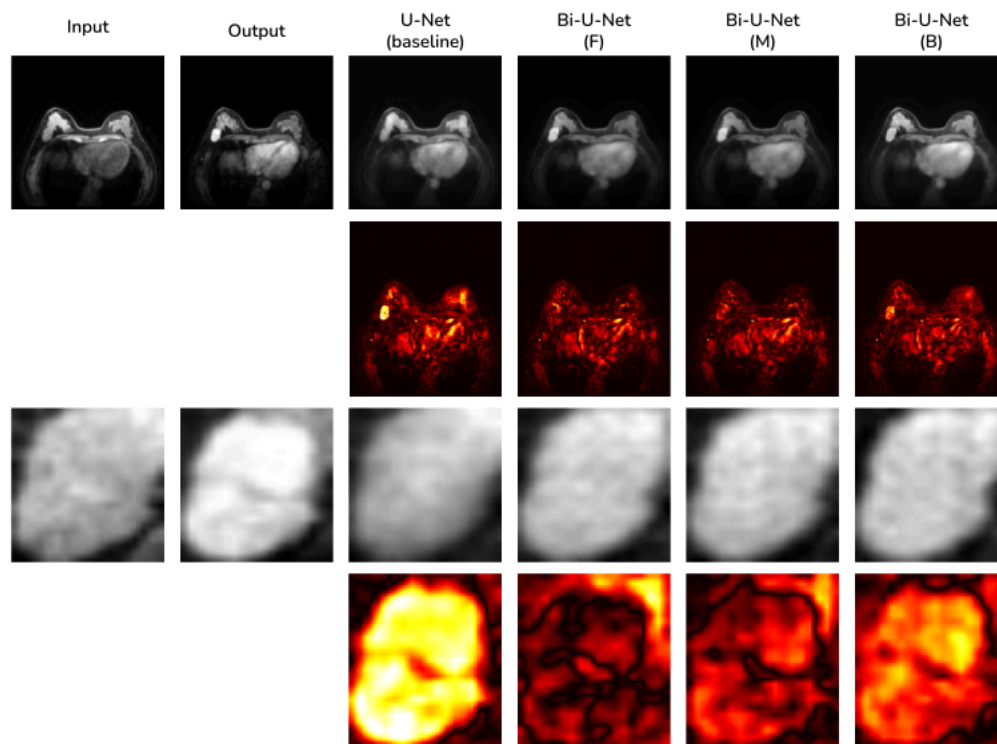


Figura 8: Comparación visual entre el modelo base (U-Net) y los modelos propuestos en este trabajo (Bi-U-Net). Fuente: Elaboración propia.

Estimación visual por los modelos Bi-GAN

Similar a la figura 8, la figura 9 muestra una comparación entre varios modelos generativos tipo GAN (**baseline**) y las tres variantes que usan el modelo Bi-U-Net como generador: Bi-GAN (F), Bi-GAN (M) y Bi-GAN (B). En este caso, la GAN (baseline) genera imágenes con una calidad aceptable, pero parece perder algunos detalles de la estructura anatómica en comparación con las variantes Bi-GAN. Similarmente al caso anterior, este modelo presenta diferencias más intensas (zonas rojas más brillantes), indicando una menor capacidad para replicar las estructuras presentes en la imagen real, particularmente en el ROI generado; en donde es menos definido, con una textura más difusa y menos capacidad para capturar detalles críticos del tumor. Sin embargo, al analizar las variantes Bi-GAN (F, M y B) Estas presentan una mejor correspondencia visual con la imagen real, logrando una mayor preservación de detalles anatómicos y contraste. En todos los casos, las variantes Bi-GAN muestran diferencias más suaves y distribuidas de manera uniforme, reflejando un desempeño superior, logrando una reducción considerable en las diferencias. Particularmente, el modelo Bi-GAN (B) obtiene el desempeño superior al mostrar las menores discrepancias, lo que confirma su superioridad en términos de fidelidad, definición de bordes, similitud estructural y claridad en general. Esto ocurre de manera similar con los ROI,

en donde Bi-GAN (B) destaca por su capacidad para mantener la forma y los detalles clave de la lesión y preservación de las características relevantes.

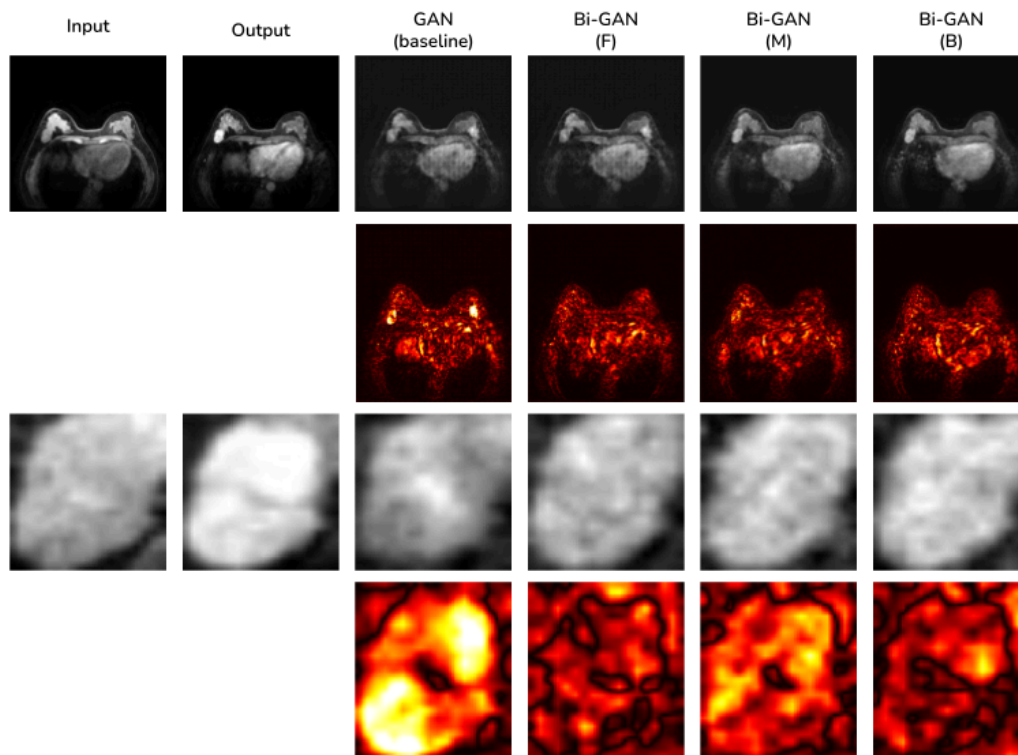


Figura 9: *Comparación visual entre el modelo base (GAN) y los modelos propuestos en este trabajo (Bi-GAN). Fuente: Elaboración propia.*

Comparación entre imágenes Bi-U-Net y Bi-GAN

Aunque ambas abordan la misma tarea de generación de imágenes post-contraste de resonancia magnética con enfoque en regiones tumorales, se observan diferencias importantes en el comportamiento de las arquitecturas.

Calidad de reconstrucción de la imagen completa: Las variantes Bi-U-Net presentan mejoras respecto al modelo base U-Net, mostrando mayor fidelidad en la reconstrucción de la imagen completa. Sin embargo, a pesar de que entre las variantes (F, M y B) son menos notorias las diferencias en escala global los modelos Bi-U-Net aún tienden a suavizar detalles, lo que genera una pérdida de precisión en las estructuras más finas. Las variantes Bi-GAN logran reconstruir imágenes con mayor calidad visual y detalles más precisos que las Bi-U-Net. En comparación con los modelos las Bi-GANs son más efectivas en preservar detalles finos en la imagen completa, lo que se refleja en una mayor similitud visual con las imágenes reales, especialmente

en regiones de alto contraste. Por otro lado, los mapas de diferencias muestran que las variantes F, M y B en general reducen considerablemente los errores en comparación con los modelos baseline. Sin embargo, la diferencia principal es encontrada en los modelos Bi-U-Net, los cuales presentan áreas con diferencias más notables, particularmente en las zonas de mayor contraste y bordes. Los mapas de diferencias de las Bi-GANs presentan errores menos intensos y mejor distribuidos que los de los demás modelos, evidenciando una capacidad superior para ajustar los detalles.

Calidad de reconstrucción de las regiones tumorales: En este caso, las variantes Bi-U-Net logran mejoras significativas en el ROI, destacando Bi-U-Net (B) por su capacidad para preservar detalles clave de la lesión. Sin embargo, la variable Bi-GAN (B) destaca especialmente por la generación de ROIs con bordes definidos y características internas más cercanas a la imagen real, lo cual evidencia una mayor fidelidad en los detalles estructurales del tumor. Esto indica que las Bi-GANs son más eficaces para captar las características finas del ROI. Por otra parte, los mapas de diferencia muestran que en general, las variantes F, M y B reducen las discrepancias, especialmente en los bordes y zonas centrales del tumor. Sin embargo, las Bi-GANs, en particular Bi-GAN (B), logran los mapas de diferencia más suaves y con menor intensidad, reflejando su capacidad para capturar con precisión las características tumorales del ROI.

6.2. Entrenamiento de los modelos de Aprendizaje

En esta sección se analiza el proceso de entrenamiento de los modelos de aprendizaje, destacando tanto las arquitecturas basadas en U-Net como en GAN. La Figura X ilustra el entrenamiento de los modelos U-Net, mientras que la Figura X presenta el entrenamiento de los modelos GAN. En ambos casos, se observa una correcta convergencia de los modelos, con la notable excepción del modelo GAN baseline, que exhibe un comportamiento inestable. Sin embargo, las variantes F, M y B muestran una notable mejora en la convergencia, logrando curvas de aprendizaje más consistentes y eficientes. En particular, la variación B se destaca al alcanzar los valores más bajos de optimización en ambos enfoques, lo que evidencia su superioridad en el proceso de aprendizaje.

Es relevante destacar que, aunque el modelo GAN baseline enfrenta inestabilidades naturales debido a la complejidad de su entrenamiento, la incorporación de las variaciones F, M y B no sólo estabiliza el proceso de aprendizaje, sino que también mejora significativamente los criterios de optimización. Esto demuestra que las modificaciones propuestas no sólo resuelven problemas de estabilidad, sino que además refuerzan la capacidad de los modelos para aprender de manera más efectiva, favoreciendo la calidad de los resultados obtenidos. Estos hallazgos validan y respaldan los análisis realizados en las secciones previas, consolidando la importancia de las variantes en el desempeño general de los modelos.

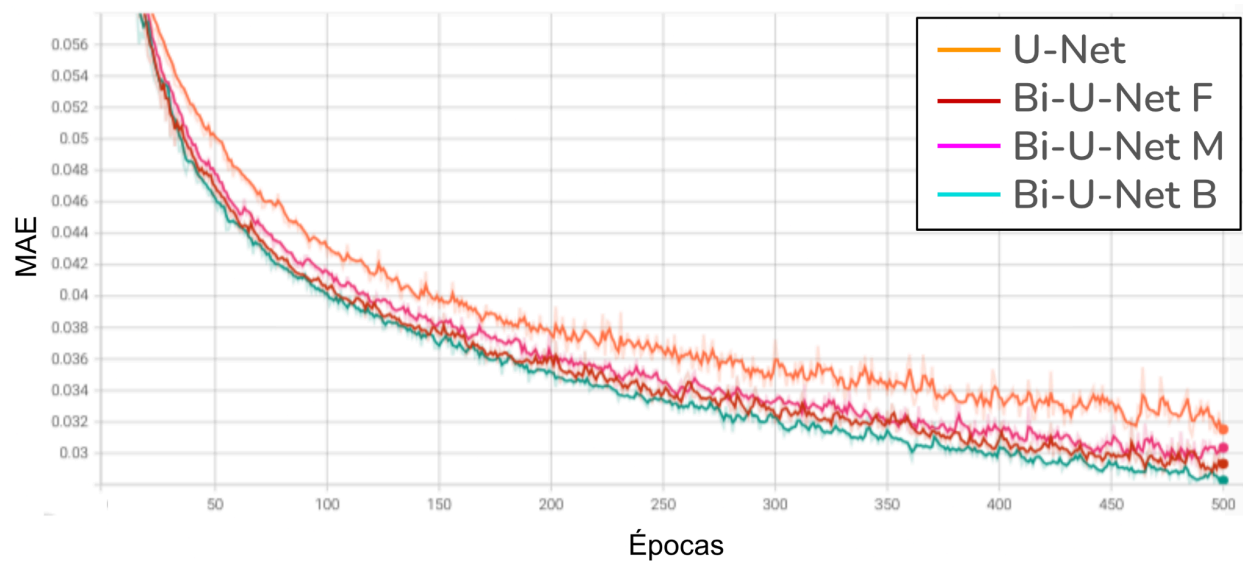


Figura 10: Loss Modelos U-Net. Fuente: Elaboración propia.

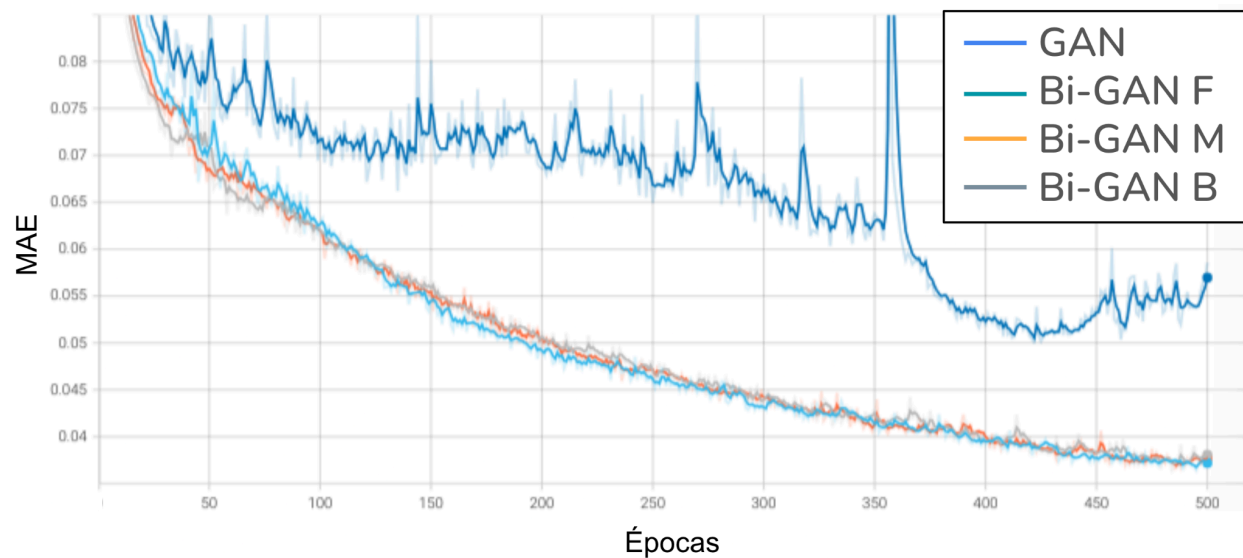


Figura 11: Loss Modelos GANs. Fuente: Elaboración propia.

7. Conclusiones

En este trabajo se presenta un modelo para la síntesis de imágenes post-contraste de MRI a partir de imágenes precontraste de mama. En el trabajo se emplean dos modelos base tipo encoder (E) y decoder (D) ampliamente usados en la síntesis de imágenes como lo son la U-Net y la GAN. Sobre estos modelos se incluye un encoder adicional (R) que extrae características de las regiones tumorales, y sobre esta modificación proponen 3 estrategias de fusión de características que permiten enfocar el aprendizaje de los modelos en las regiones tumorales. La primera involucra la fusión de los mapas de características a través de las skip connections entre los encoders y el decoder (configuración F). La segunda involucra además, la fusión de los espacios latentes extraídos a partir de E y R empleando una combinación lineal a través de una multiplicación elemento a elemento (configuración M). Por último, la tercera emplea una fusión de espacios latentes usando una combinación espacial no lineal a través de un bloque convolucional simple con activación ReLU (configuración B). Los modelos propuestos han sido nombrados como modelos Bi-U-Net (F, M y B) y modelos Bi-GAN (F, M y B). Los modelos entrenados son evaluados usando métricas como MAE, SSIM y PSNR.

Los resultados obtenidos muestran que las variaciones propuestas, tanto en las arquitecturas Bi-U-Net (F, M, B) como en las Bi-GAN (F, M, B) no solo alcanzan la precisión numérica en las métricas MAE, SSIM y PSNR, sino que también obtienen una calidad perceptual superior en las imágenes generadas en comparación con los modelos baseline. Esto es especialmente relevante en un contexto clínico, donde la preservación de estructuras anatómicas y la correcta síntesis de regiones de posible interés tumoral son esenciales para asistir en el diagnóstico médico. En particular, la variante B demuestra ser la más robusta en términos de convergencia y calidad de reconstrucción, logrando preservar estructuras relevantes y minimizar diferencias en las regiones de interés. También es importante destacar que, aunque el modelo GAN baseline presenta dificultades para alcanzar una convergencia estable debido a la complejidad de su proceso de entrenamiento, las modificaciones introducidas con todas las variantes F, M y B logran estabilizar significativamente este proceso. Esto no solo mejora los criterios de optimización, sino que también garantiza una mayor fidelidad en la generación de imágenes y una mejor correspondencia entre las regiones reales y sintetizadas. Los resultados en términos visuales y en las métricas reflejan la robustez de los modelos propuestos en la generación de imágenes post-contraste con mayor fidelidad a nivel de píxeles.

Se propone como trabajo futuro explorar la integración de modelos preentrenados con técnicas de transfer learning para potenciar la capacidad de generalización de las arquitecturas en bases de datos más amplias y diversas. Adicionalmente, sería valioso investigar la incorporación de mecanismos de atención espacial para optimizar aún más la síntesis de regiones específicas de interés, como las áreas tumorales. Por último, se pretende evaluar el impacto clínico del modelo en entornos reales mediante su validación con especialistas en radiología.

8. Referencias bibliográficas

- [1] González, F. A. (2015). Modelos de aprendizaje computacional en reumatología. *Revista Colombiana de Reumatología*, 22(2), 77-78.
- [2] Osorio BN, Bello HC, Vega BL. Factores de riesgo asociados al cáncer de mama. *Rev Cubana Med Gen Integr.* (2020); 36(2):1-13.
- [3] Siegel RL , Giaquinto AN , Jemal A . Cancer statistics, 2024. *CA Cancer J Clin.* 2024; 74(1): 12-49. doi:10.3322/caac.21820
- [4] Mann, R. M., Kuhl, C. K., & Moy, L. (2019). Contrast-enhanced MRI for breast cancer screening. *Journal of Magnetic Resonance Imaging*, 50(2), 377-390.
- [5] Debelee, T.G., Schwenker, F., Ibenhal, A. et al. Survey of deep learning in breast cancer image analysis. *Evolving Systems* 11, 143–163 (2020). <https://doi.org/10.1007/s12530-019-09297-2>
- [6] Ruiz Baquerizo, K. R. (2021). Contribución de la implementación de Inteligencia Artificial para el diagnóstico precoz del cáncer de mama (Bachelor 's thesis, Quito: UCE)./http://www.dspace.uce.edu.ec/handle/25000/24990
- [7] García, M., & López, J. (2023). La disponibilidad de dataset de resonancia magnética post-contraste etiquetados y de alta calidad. *Revista de Tecnología Médica*, 15(2), 123-134.
- [8] Navarro, M.E.; Razmilic, D.; Araos, I.; Rodrigo, A.; Andía, M.E. Rendimiento de la mamografía espectral de energía dual con contraste en la detección de cáncer de mama: experiencia en un centro de referencia. *Rev Med Chile* 2018, 146(2), 141-149. https://www.scielo.cl/scielo.php?pid=S0034-98872018000200141&script=sci_arttext&tlng=en.
- [9] Romo-Sanchez, M.; Nelson, F.; Sangrador-Deitos, M. Magnetic Resonance Imaging in Multiple Sclerosis: a review of the basic principles and practical guidelines. *Arch Neurociencias* 2021, 25(4), 23-31. <https://www.medigraphic.com/cgi-bin/new/resumenI.cgi?IDARTICULO=98775>.
- [10] Zhao, X.; Liao, Y.; Xie, J.; He, X.; Zhang, S.; Wang, G.; Fang, J.; Lu, H.; Yu, J. BreastDM: A DCE-MRI dataset for breast tumor image segmentation and classification. *Comput Biol Med* 2023, 164(107255), 107255. <https://doi.org/10.1016/j.compbio.2023.107255>.
- [11] Quillent, A.; Bismuth, V.; Bloch, I.; Kervazo, C.; Ladjal, S. A deep learning method trained on synthetic data for digital breast tomosynthesis reconstruction (s/f). *GE HealthCare*.
- [12] Moreno, D.U.; Suárez, J.C.U.; Rondón, B.D.; Rodríguez, A.N.; Peña, R.T.; de la Nuez, J.G. Resonancia magnética en el diagnóstico de las lesiones sólidas mamarias. *Investig Medicoquirúrgicas* 2018, 10(2). <https://www.medigraphic.com/cgi-bin/new/resumen.cgi?IDARTICULO=85530>.
- [13] García Pérez, J.C. Detección y segmentación automática de lesiones en pacientes con esclerosis múltiple en imágenes de resonancia magnética. <https://openaccess.uoc.edu/webapps/o2/handle/10609/99167>.
- [14] Escalante González, M. Aplicación de la inteligencia artificial para la detección del cáncer de mama. *Rev Med Sinergia* 2023, 8(12), e1113. <https://doi.org/10.31434/rms.v8i12.1113>.
- [15] García, M.; López, J. La disponibilidad de dataset de resonancia magnética post-contraste etiquetados y de alta calidad. *Rev Tecnol Med* 2023, 15(2), 123-134.
- [16] (N.d.). Unirioja.Es. La disponibilidad de dataset de resonancia magnética post-contraste etiquetados y de alta calidad. Recuperado el 21 de octubre de 2024, de <https://dialnet.unirioja.es/servlet/articulo?codigo=9430142>.

- [17] (N.d.). Edu.Co. Retrieved September 3, 2024, from <https://repositorio.itm.edu.co/handle/20.500.12622/6495>
- [18] Zhao, X., Liao, Y., Xie, J., He, X., Zhang, S., Wang, G., Fang, J., Lu, H., & Yu, J. (2023). BreastDM: A DCE-MRI dataset for breast tumor image segmentation and classification. *Computers in Biology and Medicine*, 164(107255), 107255. <https://doi.org/10.1016/j.compbiomed.2023.107255>
- [19] (N.d.). Amazon.com. Retrieved November 4, 2024, from [https://aws.amazon.com/es/what-is/gan/#:~:text=Una%20red%20generativa%20antag%C3%B3nica%20\(GAN,de%20datos%20de%20entrenamiento%20determinado.](https://aws.amazon.com/es/what-is/gan/#:~:text=Una%20red%20generativa%20antag%C3%B3nica%20(GAN,de%20datos%20de%20entrenamiento%20determinado.)
- [20] Daniel. (2022, April 26). U-NET: todo lo que tienes que saber sobre la red neuronal de Computer Vision. Formación en ciencia de datos | DataScientest.com; DataScientest. <https://datascientest.com/es/u-net-lo-que-tienes-que-saber>
- [21] Zhao, X., Liao, Y., Xie, J., He, X., Zhang, S., Wang, G., ... & Yu, J. (2023). BreastDM: A DCE-MRI dataset for breast tumor image segmentation and classification. *Computers in Biology and Medicine*, 164, 107255.
- [22] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18* (pp. 234-241). Springer International Publishing.
- [23] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139-144.
- [24] Adiyaman, H., Emre Varul, Y., Bakırman, T., & Bayram, B. (2024). Stripe Error Correction for Landsat-7 Using Deep Learning. *PFG—Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 1-13.
- [25] Alam, F. F., Rahman, T., & Shehu, A. (2020). Evaluating autoencoder-based featurization and supervised learning for protein decoy selection. *Molecules*, 25(5), 1146. <https://doi.org/10.3390/molecules25051146>
- [26] Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786), 504-507. <https://doi.org/10.1126/science.1127647>
- [27] Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8), 1798-1828. <https://doi.org/10.1109/TPAMI.2013.50>
- [28] McRobbie, D. W., Moore, E. A., & Graves, M. J. (2007). *MRI from picture to proton*. Cambridge University Press.
- [29] Kuhl, C. K., & Bickel, H. (2005). Dynamic contrast-enhanced MRI in breast cancer: A review of the literature and clinical applications. *European Journal of Radiology*, 56(1), 1-10. <https://doi.org/10.1016/j.ejrad.2005.02.010>
- [30] Tofts, P. S., & Kermode, A. G. (1991). Measurement of the blood-brain barrier permeability and leakage space using dynamic MRI. *Magnetic Resonance in Medicine*, 17(2), 357-367. <https://doi.org/10.1002/mrm.1910170212>
- [31] Chavhan, G. B., & Babyn, P. S. (2010). Contrast-enhanced magnetic resonance imaging: Principles and applications in pediatric imaging. *Pediatric Radiology*, 40(11), 1844-1854. <https://doi.org/10.1007/s00247-010-1820-2>