

**GENERACIÓN DE SECUENCIAS CINE-MRI CARDÍACAS UTILIZANDO
APRENDIZAJE GENERATIVO ADVERSARIO CONDICIONADO POR
INFORMACIÓN DEL CORAZÓN**

HENRY IVÁN PEÑA CONTRERAS

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

2021

**GENERACIÓN DE SECUENCIAS CINE-MRI CARDÍACAS UTILIZANDO
APRENDIZAJE GENERATIVO ADVERSARIO CONDICIONADO POR
INFORMACIÓN DEL CORAZÓN**

HENRY IVÁN PEÑA CONTRERAS

**Trabajo de investigación presentado para optar al título de Ingeniero de
Sistemas**

Director:

Fabio Martínez Carrillo

Ph.D en Ingeniería de Sistemas y Computación

Co-Director:

David Edmundo Romo Bucheli

Ph.D en Ingeniería Eléctrica

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

2021

El autor expresa su agradecimiento:

Al grupo de investigación Biomedical Imaging, Vision and Learning Laboratory (Biv-L²ab) y las personas que me acompañaron durante toda la carrera, ya que el vínculo creado con estas personas fue similar al de una familia; en especial al profesor y mi director de tesis Fabio Martínez Carrillo quien, además de formarme en el campo de la investigación, también me ayudó a madurar como persona, y por ende me encuentro muy agradecido. Por otro lado, quiero expresar mi agradecimiento a mis padres; a mi padre quien, a pesar de ya no tenerlo a mi lado, me dio la mejor formación e infancia que pudiera haber deseado, además de que me introdujo en el campo de lo que será mi futuro entorno laboral. También, quisiera agradecer a mi madre, por su paciencia, apoyo moral y afecto cuando más lo necesité, ya que gracias a estas fue más llevadero enfrentar los diferentes retos y adversidades.

A Saray, que ha sido, y es, una de las personas más importantes que he conocido y con la que he estado en estos últimos momentos de carrera, tanto académica como sentimentalmente. También quisiera agradecer a mis amigos más cercanos, Oscar, Nicolás, Edgar, Douglas, Andrés, Diego, por su amistad y apoyo, y a mi familia, en especial a mi primo Luis, que en los últimos momentos de mi tesis fue un gran apoyo.

A la escuela de Ingeniería de Sistemas e Informática (EISI) y todos sus docentes que me guiaron y formaron para ser un excelente profesional. Al Ingeniero Santiago Gómez, quien me guió y enseñó mucho sobre el tema de mi tesis, y me ayudó en todo lo que pudo, por lo que me encuentro muy agradecido.

Y por último, pero no menos importante, a Dios, que a pesar de no ser muy fiel a creencias, considero que me ayudó en muchos momentos de incertidumbre.

CONTENIDO

| | pág. |
|---|-----------|
| INTRODUCCIÓN | 11 |
| 1. MARCO TEÓRICO Y TRABAJOS PREVIOS | 16 |
| 1.1. REDES GENERATIVAS ADVERSARIAS (GANs) | 17 |
| 1.2. BIOMARCADORES, VECTORES EMBEBIDOS Y REPRESENTACIONES | 23 |
| 1.3. DE PATRONES CARDIACOS A REPRESENTACIONES PROFUNDAS | 26 |
| 2. PLANTEAMIENTO Y JUSTIFICACIÓN DEL PROBLEMA | 29 |
| 3. OBJETIVOS | 30 |
| 4. METODOLOGÍA PROPUESTA | 31 |
| 4.1. ESQUEMA GENERATIVO CON SOPORTE A MÚLTIPLES CONDICIONES | 32 |
| 4.1.1. CONSISTENCIA EN LAS CARACTERÍSTICAS | 35 |
| 4.2. CONDICIONES CARDIACAS A PRIORI | 36 |
| 4.2.1. SEGMENTACIÓN DEL CORAZÓN | 36 |
| 4.2.2. FLUJO ÓPTICO | 39 |
| 4.2.3. PATOLOGÍA ASOCIADA | 43 |
| 4.2.4. GENERACIÓN DE EMBEBIDOS | 45 |
| 4.2.5. AGRUPAMIENTO DESDE UMAP | 45 |
| 4.2.6. ESTRATEGIAS DE CLASIFICACIÓN | 47 |
| 4.2.7. K-VECINOS CERCANOS | 47 |
| 4.2.8. MÁQUINAS DE SOPORTE VECTORIAL | 48 |
| 4.2.9. BOSQUE ALEATORIO | 50 |
| 4.3. DISEÑO EXPERIMENTAL | 52 |

| | |
|--|-----------|
| 4.3.1. SELECCIÓN DE DATOS | 52 |
| 4.3.2. CONFIGURACIÓN DE LA MCIGAN | 53 |
| 4.3.3. VECTORES EN ESPACIO LATENTE | 54 |
| 4.3.4. VALIDACIÓN ESTADÍSTICA | 54 |
| 5. EVALUACIÓN Y RESULTADOS | 56 |
| 5.1. REPRESENTACIÓN PROFUNDA CONDICIONADA CON INFORMACIÓN MORFOLÓGICA: SEGMENTACIÓN | 58 |
| 5.2. REPRESENTACIÓN PROFUNDA CONDICIONADA CON INFORMACIÓN DE MOVIMIENTO: FLUJO ÓPTICO | 61 |
| 5.3. REPRESENTACIÓN PROFUNDA CONDICIONADA CON INFORMACIÓN MORFOFISIOLÓGICA INTEGRADA | 64 |
| 6. CONCLUSIONES Y PERSPECTIVAS | 67 |
| BIBLIOGRAFÍA | 70 |

LISTA DE FIGURAS

| | pág. |
|--|-------------|
| Figura 1. Arquitectura de una GAN convolucional | 18 |
| Figura 2. Cuadro comparativo CycleGAN vs MR-GAN | 22 |
| Figura 3. <i>Pipeline</i> del trabajo sobre desarrollo de biomarcadores por medio de redes neuronales | 25 |
| Figura 4. Segmentación del corazón | 27 |
| Figura 5. Esquema de red neuronal para segmentar | 28 |
| Figura 6. <i>Pipeline</i> de la metodología propuesta. | 32 |
| Figura 7. Arquitectura de la red generativa adversaria condicional propuesta. | 33 |
| Figura 8. Arquitectura U-NET usada para la segmentación de resonancias magnéticas cardiacas. | 38 |
| Figura 9. Ejemplo de una segmentación aprendida por el modelo U-NET. . . | 39 |
| Figura 10. Ejemplo de flujo óptico en una resonancia magnética cardiaca. . . | 42 |
| Figura 11. Ejemplo de flujo óptico segmentado en una resonancia magnética cardiaca. | 43 |
| Figura 12. Metodología para la obtención de embebidos | 46 |
| Figura 13. Ejemplo del algoritmo KNN | 48 |
| Figura 14. Ejemplo de hiperplano de separación en un espacio bidimensional. | 49 |
| Figura 15. Ejemplo de clasificación por medio de un árbol de decisión junto con su diagrama. | 51 |
| Figura 16. Ejemplo de clasificación para distintos modelos de aprendizaje. . . | 52 |
| Figura 17. Generaciones de la arquitectura con los diferentes prioris a evaluar. | 57 |

| | |
|--|----|
| Figura 18. Agrupamiento no supervisado UMAP de los vectores embebidos condicionados con la segmentación. | 59 |
| Figura 19. Matrices de confusión con Segmentación | 59 |
| Figura 20. métricas segmentación | 60 |
| Figura 21. Agrupamiento no supervisado UMAP de los vectores embebidos condicionados con el flujo. | 62 |
| Figura 22. Matrices de confusión con Flujo óptico | 63 |
| Figura 23. métricas flujo | 64 |
| Figura 24. Agrupamiento no supervisado UMAP de los vectores embebidos condicionados con la segmentación + flujo. | 65 |
| Figura 25. Matrices de confusión con Flujo óptico en el área segmentada . . . | 65 |
| Figura 26. métricas flujo óptico segmentado | 66 |

LISTA DE CUADROS

| | pág. |
|---|-------------|
| Cuadro 1. Hiperparámetros de entrenamiento para el modelo de Knn | 55 |
| Cuadro 2. Hiperparámetros de entrenamiento para el modelo de Random Forest | 55 |
| Cuadro 3. Hiperparámetros de entrenamiento para el modelo de support vec- tor machine. | 55 |

RESUMEN

TÍTULO: GENERACIÓN DE SECUENCIAS CINE-MRI CARDÍACAS UTILIZANDO APRENDIZAJE GENERATIVO ADVERSARIO CONDICIONADO POR INFORMACIÓN DEL CORAZÓN *

AUTOR: HENRY IVÁN PEÑA CONTRERAS **

PALABRAS CLAVE: MODELOS GENERATIVOS, REDES GENERATIVAS ADVERSARIAS CONDICIONALES, RESONANCIA MAGNÉTICA, CICLOS CARDÍACOS.

DESCRIPCIÓN: Las enfermedades cardiacas son el mayor motivo de muerte tanto a nivel local como global, registrando mas de 200.000 muertes cada año. Las secuencias cine-MRI soportan la detección temprana y seguimiento de patologías cardiacas, permitiendo la valoración morfológica y fisiológica de las paredes del corazón durante el ciclo cardiaco. Esta valoración, sin embargo, es hoy en día altamente dependiente del experto observador, lo que puede conducir a una alta subjetividad en su análisis. Técnicas de análisis de imágenes y algoritmos de aprendizaje de máquina, surgen entonces, como herramientas alternativas para soportar el diagnóstico clínico, reducir la variabilidad en los análisis y brindar nuevas descripciones y correlaciones de los patrones que influyen en el buen funcionamiento cardiaco. Este trabajo desarrolló un modelo computacional basado en redes generativas adversarias condicionales para la generación de secuencias cardiacas. Además, se estudiaron mecanismos *a priori* que permitan generar secuencias sintéticas de cine-MRI coherente y con sentido de ciertas patologías de entrenamiento. La generación de estas secuencias, guiadas por un patrón *a priori* permitió generar descriptores embebidos en la arquitectura, los cuales fueron utilizados como nuevos marcadores de la enfermedad frente a diferentes modelos de clasificación, con el fin de describir patrones espacio-temporales aprendidos desde un conjunto de secuencias de entrenamiento.

* Trabajo de investigación

** Facultad de Ingenierías Físico-Mecánicas. Escuela de Ingeniería de Sistemas e Informática. Director: Fabio Martínez Carrillo, Ph.D. en Ingeniería de Sistemas y Computación. Codirector: David Edmundo Romo Bucheli, Ph.D en Ingeniería Eléctrica.

ABSTRACT

TITLE: CARDIAC CINE-MRI SEQUENCES SYNTHESIS USING GENERATIVE ADVERSARIAL LEARNING CONDITIONED BY HEART INFORMATION. *

AUTHOR: HENRY IVÁN PEÑA CONTRERAS **

KEYWORDS: GENERATIVE MODEL, CONDITIONAL GENERATIVE ADVERSARIAL NETWORK, MRI-SEQUENCES, CARDIAC CYCLES.

DESCRIPTION: Heart disease is the largest single cause of death both locally and globally, accounting for more than 200,000 deaths each year. MRI-cinema sequences support the early detection and monitoring of cardiac pathologies, allowing the morphological and physiological assessment of the heart walls during the cardiac cycle. This evaluation, however, is nowadays highly dependent on the expert observer, which can lead to a high subjectivity in its analysis. Image analysis techniques and machine-learning algorithms then emerge as alternative tools to support clinical diagnosis, reduce variability in analysis, and provide new descriptions and correlations of the patterns that influence good cardiac function. This work developed a computational model based on conditional adverse generation networks for the generation of cardiac sequences. In addition, mechanisms *a priori* that allow the generation of synthetic sequences of coherent and meaningful MRI-cinema of certain training pathologies were studied. The generation of these sequences, guided by an *a priori* pattern, allowed the generation of descriptors embedded in the architecture, which were used as new disease markers against different classification models, in order to describe spatial-temporal patterns learned from a set of training sequences.

* Research work

** Faculty of Physical-Mechanical Engineering. School of Systems and Computer Engineering. Advisor: Fabio Martínez Carrillo, Ph.D. in Computer and Systems Engineering. Co-Advisor: David Edmundo Romo Bucheli, Ph.D in Electrical Engineering

INTRODUCCIÓN

Las enfermedades cardiovasculares son la principal causa de muerte a nivel mundial. De hecho, la Organización Mundial de la Salud estima que en el 2030 más de 23,6 millones de personas morirán por alguna enfermedad cardiovascular¹. Sumado a lo anterior y en relación a la pandemia que afrontamos actualmente, se ha demostrado que pacientes con deficiencias miocárdicas, infectados con el virus COVID-19, son más vulnerables a desarrollar el síndrome de dificultad respiratoria aguda.². El diagnóstico correcto de enfermedades cardíacas, su respectivo seguimiento y valoración es fundamental para mitigar y afrontar estas afecciones. Las resonancias magnéticas cardíacas (cines MRI) representan una modalidad de imágenes diagnósticas que permiten analizar la anatomía cardíaca y su función ventricular. Estas secuencias por lo tanto se convierten en una de las herramientas más importantes para el soporte durante la rutina clínica. El análisis y valoración de las afecciones cardíacas, sobre estas secuencias cine-MRI, es llevada a cabo por cardiólogos basada casi exclusivamente en su experiencia. Este hecho resulta altamente subjetivo y propenso a errores. En valoraciones cuantitativas, los análisis sobre los cine-MRI se limitan a la cuantificación de patrones hemodinámicos y reglas determinísticas que pueden ser sensibles a la variabilidad de las diferentes enfermedades cardíacas existentes. Además, el cálculo de estos patrones hemodinámicos son dependientes de una apropiada segmentación de los ventrículos en la fase del final de la diástole y el final de la sístole. Estas segmentaciones son comúnmente llevadas a cabo de

¹ MENDIS, Shanthi, *et al.* *Global atlas on cardiovascular disease prevention and control*. World Health Organization, 2011.

² BONOW RO Fonarow GC, O'Gara PT Yancy CW. *Association of Coronavirus Disease 2019 (COVID-19) With Myocardial Injury and Mortality*. 2020.

forma manual, siendo dispendiosas, inclusive para expertos cardiólogos.

Las herramientas computacionales y sistemas de soporte al diagnóstico (CAD's por sus siglas en inglés) han permitido mitigar estas limitaciones en las valoraciones cardíacas y soportar los análisis llevados durante la rutina clínica. Entre otras, estas estrategias han permitido disminuir la subjetividad, guiar los procedimientos clínicos y determinar nuevas relaciones de variables cardíacas correlacionadas con patologías específicas ³. Dentro de las tareas más comunes de estos sistemas, se puede mencionar, la segmentación de los ventrículos, la cuantificación de patrones cardiovasculares, la predicción de patologías, entre otros. Hoy en día, muchos de los algoritmos que componen estos sistemas están basados en estrategias de aprendizaje de máquina, siendo dependientes de un amplio número de anotaciones dadas por expertos. La asignación de etiquetas por profesionales con diferente grado de experticia puede conducir a que el comportamiento de estos algoritmos sea bastante limitado y se requiera una supervisión o ajuste final de un experto. Además, estos algoritmos en términos supervisados, se limitan a imitar las tareas típicas de los cardiólogos o identificar patrones comúnmente conocidos en el área.

Con el desarrollo de nuevas y cada vez más robustas representaciones de aprendizaje de máquina se ha podido impactar en nuevos y más sofisticados métodos para la ingeniería biomédica. Esto ha conllevado, en diferentes dominios, a buscar nuevos descriptores y biomarcadores digitales que puedan brindar nuevas relaciones estructurales y dinámicas que puedan estar asociadas a patologías particulares. Los algoritmos generativos-adversarios han ganado recientemente gran importancia para generar nuevas imágenes coherentes a partir de vectores ruidosos, teniendo durante el proceso diferentes niveles de descripción de las imágenes. Además, en niveles compactos de esta estructura, se pueden recuperar vectores embebidos

³ KONONENKO, Igor. "Machine learning for medical diagnosis: history, state of the art and perspective". En: *Artificial Intelligence in medicine* 23.1 (2001), págs. 89-109.

("embeddings") que describen relaciones no lineales, que permiten la apropiada formación y descripción de la imagen ⁴. Por ejemplo, en el estado del arte existen arquitecturas y trabajos enfocados en la generación de secuencias cardíacas los cuales han generado resultados destacables relacionados con la capacidad de generar estructuras con coherencia cardíaca. Particularmente en el trabajo de Zhang et al⁵, se utiliza una arquitectura *CycleGAN* para producir secuencias de cine-MRI utilizando consistencia de forma desde secuencias tomográficas. Esta estrategia reporta algunas limitaciones para sintetizar secuencias con coherencia morfológica. También, la arquitectura llamada *SCGANS*, genera resonancias teniendo en cuenta la correlación de información cardíaca, disponible en tomografías asociadas ⁶. Algunas limitaciones de esta estrategia son parches generados con una calidad deficiente, la cual relaciona esto a un problema para emular las estructuras cardíacas, con o sin una patología presente. Estos trabajos han desarrollado representaciones generativas, pero enfocadas en aumentar datos cine-MRI para resolver tareas de delineación y segmentación de los ventrículos. Sin embargo, estas representaciones no han sido valoradas para el análisis de diferentes patrones cardíacos con relación a ciertas patologías. En otros dominios estas arquitecturas han permitido, por ejemplo, la detección de tumores metastásicos en los huesos utilizando un paradigma no supervisado al entrenar una arquitectura *ANOGAN* con tomografías computarizadas donde hay tumores óseos no metastásicos, logrando un óptimo reconocimiento

⁴ LI, Ziqiang, et al. *Interpreting the Latent Space of GANs via Correlation Analysis for Controllable Concept Manipulation*. 2020. arXiv: 2006.10132 [cs.CV].

⁵ WONG, Sebastien C, et al. "Understanding data augmentation for classification: when to warp?" En: *2016 international conference on digital image computing: techniques and applications (DICTA)*. IEEE. 2016, págs. 1-6.

⁶ PEREZ, Luis y WANG, Jason. "The effectiveness of data augmentation in image classification using deep learning". En: *arXiv preprint arXiv:1712.04621* (2017).

de estos.⁷ También esta arquitectura se ha utilizado en problemas como el de clasificación de cáncer de mama⁸, la cual provee nuevos biomarcadores y ayuda a reducir la carga de trabajo en la observación manual de los patólogos.

En este trabajo se desarrolló una estrategia basada en algoritmos generativos-adversarios para la construcción de una representación profunda cardíaca, que pueda ser explotada a diferentes niveles jerárquicos y permita evaluar posibles correlaciones con patologías. Una contribución de este trabajo fue la introducción y evaluación de diferentes fuentes de información *a priori* como condicionadores de la representación cardíaca con sentido clínico y relación con la estructura de las paredes cardíacas, o información relacionada con su dinámica. En este trabajo se utilizó una arquitectura de múltiples entradas y se condicionó información relacionada con la segmentación del ventrículo izquierdo, así como también, máscaras de flujo aparente y la patología asociada. Para el modelamiento de la segmentación del ventrículo se utilizó una red U-net y para el flujo aparente se tuvieron en cuenta desplazamientos largos, calculados desde el algoritmo de Brox, quién realiza una optimización de diversas restricciones con el objetivo de obtener la velocidad relativa del corazón a través del ciclo cardiaco. Una vez entrenado el modelo la parte del generador fue adaptado como la representación profunda cardiaca, utilizando una retro-propagación de las imágenes generadas hacia un vector embebido que representaba la enfermedad. Los vectores embebidos fueron mapeados a un espacio de baja dimensionalidad, usando una estrategia de UMAP y posteriormente clasificados con diferentes modelos de clasificación según la condición patológica.

⁷ WATANABE, H., *et al.* "Bone Metastatic Tumor Detection based on AnoGAN Using CT Images". En: *2019 IEEE 1st Global Conference on Life Sciences and Technologies (LifeTech)*. 2019, págs. 235-236. DOI: 10.1109/LifeTech.2019.8883999.

⁸ MAN, R.; YANG, P. y XU, B. "Classification of Breast Cancer Histopathological Images Using Discriminative Patches Screened by Generative Adversarial Networks". En: *IEEE Access* 8 (2020), págs. 155362-155377. DOI: 10.1109/ACCESS.2020.3019327.

Los resultados demuestran consistencia en la generación de secuencias cine-MRI, mostrando la mejor configuración de condiciones el experimento de flujo óptico localizado en la segmentación junto con la patología asociada, logrando obtener un 90,17% en la precisión y sensibilidad de un 89,95%. En los siguientes capítulos se describe en detalle el problema y los objetivos planteados, así como también, el método propuesto y la evaluación obtenida. Al final del documento se discute y se presentan algunas conclusiones del trabajo realizado.

1. MARCO TEÓRICO Y TRABAJOS PREVIOS

Los modelos de aprendizaje de máquina, en especial los modelos de aprendizaje profundo han sido recientemente utilizados en aplicaciones médicas debido a su demostrada capacidad para soportar tareas como la clasificación, segmentación, detección, súper resolución, reconstrucción, entre muchos otros ⁹. Estos modelos han mostrado resultados efectivos debido en gran medida a la cantidad y diversidad de datos permitiendo resolver nuevas y más desafiantes tareas ¹⁰. Estos modelos también han sido adaptados y usados como soporte en investigaciones que buscan encontrar nuevos patrones, biomarcadores y relaciones que permitan explicar y soportar condiciones malignas. Para ello se han utilizado modelos de aprendizaje profundo, los cuales permiten obtener representaciones jerárquicas, con capas intermedias y ocultas, según un criterio *a priori*. Estos descriptores aprendidos pueden ser usados como alternativa para evidenciar nuevos patrones asociados a patologías particulares. Teniendo en cuenta el principal objetivo de esta propuesta, a continuación se describe los fundamentos y trabajos relacionados con generación de secuencias cardíacas, la construcción de representaciones profundas generativas, los vectores embebidos y aplicaciones similares reportadas en el estado del arte.

⁹ ZHOU, S Kevin; GREENSPAN, Hayit y SHEN, Dinggang. *Deep learning for medical image analysis*. Academic Press, 2017.

¹⁰ YI, Xin; WALIA, Ekta y BABYN, Paul. "Generative adversarial network in medical imaging: a review". En: *arXiv preprint arXiv:1809.07294* (2018).

1.1. REDES GENERATIVAS ADVERSARIAS (GANs)

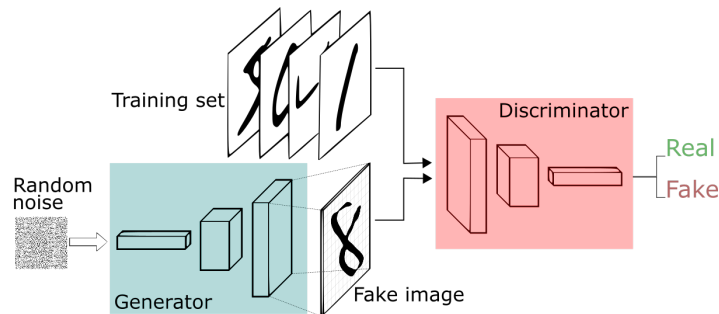
Las redes convolucionales profundas constituyen el estado del arte para resolver un gran número de tareas en visión por computador y tareas relacionadas con análisis de imágenes. En resumen, estas redes construyen representaciones multi-nivel y aprenden representaciones no lineales en cada capa, que en conjunto permiten responder a tareas específicas supervisadas. Dependiendo las tareas de entrenamiento, los datos disponibles y las salidas generadas, en la literatura se han propuesto diferentes arquitecturas convolucionales ¹¹.

Las redes generativas adversarias (GAN) han surgido como una alternativa sofisticada para generar ejemplos sin tener en cuenta el modelamiento de una distribución particular. Una principal ventaja de este hecho resulta en el libre aprendizaje de correlaciones de las observaciones con las instancias de entrenamiento, que pueden ser primordiales para detectar eventos escasos pero con alta correlación a cierta instancia de entrenamiento. En particular, una red generativa adversaria (GAN) consta de dos redes neuronales: La primera, un generador G cuyo objetivo es generar imágenes que se parezcan a un conjunto de datos con el que se entrena. La segunda, un discriminador D cuya tarea es diferenciar entre los grupos de imágenes reales y generadas, que permita guiar a G a producir mejores imágenes. Estos modelos generativos permiten producir nuevas muestras de una distribución de datos deseada, sin modelar explícitamente la función densidad de probabilidad de los datos ¹². En la figura 1 se presenta el esquema general de funcionamiento de una GAN.

¹¹ ABRAHAM, Gilu K.; JAYANTHI, V.S. y BHASKARAN, Preethi. "10 - Convolutional neural network for biomedical applications". En: *Computational Intelligence and Its Applications in Healthcare*. Ed. por VERMA, Jitendra Kumar; PAUL, Sudip y JOHRI, Prashant. Academic Press, 2020, págs. 145 -156. DOI: <https://doi.org/10.1016/B978-0-12-820604-1.00010-8>.

¹² GOODFELLOW, Ian, *et al.* "Generative adversarial nets". En: *Advances in neural information processing systems*. 2014, págs. 2672-2680.

Figura 1. Arquitectura de una GAN convolucional



Fuente: Palash SHARMA. *Autoencoders in Keras – Introduction to Beginners with Example*. URL: <https://machinelearningknowledge.ai/autoencoders-in-keras-introduction-to-beginners-with-example/>

Estos esquemas GANs han logrado resultados importantes tanto en la comunidad científica, como en la industria, permitiendo la generación de imágenes de alta calidad y el modelamiento de dominios complejos en escenarios con información de fondo redundante. No obstante, estas redes generativas adversarias, aún presentan múltiples desafíos para ser reproducidas en áreas particulares, por ejemplo, deben ser entrenadas durante largos periodos de tiempo y tienen problemas de convergencia a la distribución deseada ¹³.

Particularmente, en el dominio de imágenes médicas, las redes generativas (GANs) resultan fundamentales para la síntesis de secuencias, teniendo en cuenta protocolos específicos de adquisición, variaciones naturales de tamaño, forma y localización, de diferentes órganos frente a diferentes patologías. Una motivación principal es la generación de secuencias médicas que permitan explorar y analizar estructu-

¹³ GOODFELLOW, Ian. "NIPS 2016 tutorial: Generative adversarial networks". En: *arXiv preprint arXiv:1701.00160* (2016).

ras particulares de interés, de esta manera, se mantiene la privacidad de los pacientes y se aumenta la cantidad de ejemplos que logren modelar patologías específicas con un banco de datos mínimo. Más importante aún, las GANs pueden brindar apoyo y ayudar al desarrollo de herramientas para soporte al diagnóstico. En este sentido, el componente discriminatorio D puede también entenderse y modelarse como un regularizador, con la capacidad de identificar anomalías en las secuencias clínicas, y por lo tanto, identificar patologías a partir de la morfología registrada en las secuencias.

En la actualidad se han propuesto diferentes modelos y aplicaciones de GANs en secuencias de resonancias magnéticas para el aumento de datos, siendo libre de términos condicionales o introduciendo restricciones según regularizadores como texto e imágenes ¹⁰. Francesco Calimeri ¹⁴ propuso una red GAN no condicional, LAPGAN, para generar MRI del cerebro y un validador convolucional, entrenado previamente. Como también, se ha propuesto una red condicional con información previa de regiones de interés para un mejor rendimiento (Recon-GLGAN) ¹⁵. También se realizó una comparación¹⁶ entre la arquitectura (DCGAN) ¹⁷ y la arquitectura

¹⁴ CALIMERI, Francesco , *et al.* "Biomedical data augmentation using generative adversarial neural networks". En: *International conference on artificial neural networks*. Springer. 2017, págs. 626-634.

¹⁵ BALAMURALI MURUGESAN Vijaya Raghavan S, Kaushik Sarveswaran Keerthi Ram Mohanasankar Sivaprakasam. "A Global-Local context based Generative Adversarial Network for MRI Reconstruction". En: *arXiv preprint arXiv:1908.09262* (2019).

¹⁶ HAN, Changhee, *et al.* "GAN-based synthetic brain MR image generation". En: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE. 2018, págs. 734-738.

¹⁷ RADFORD, Alec; METZ, Luke y CHINTALA, Soumith. "Unsupervised representation learning with deep convolutional generative adversarial networks". En: *arXiv preprint arXiv:1511.06434* (2015).

Wasserstein GAN (WGAN) ¹⁸, en cuanto a la capacidad para producir imágenes cerebrales fiables, de alta resolución, mostrando resultados más realistas en la arquitectura WGAN. Estos favorables resultados pueden ser justificados debido a la estabilidad de la WGAN en el entrenamiento gracias a su función de costo, la cual lleva al modelo a mejorar su habilidad para generalizar y producir una cantidad mayor de ejemplos. En el estado del arte también se utiliza la arquitectura DCGAN¹⁹, ilustrando algunas limitaciones para reproducir estructuras anatómicas y de luminosidad en imágenes cerebrales. Incluso se han obtenido volúmenes cerebrales completos en 3D a partir de un pequeño conjunto de datos de entrenamiento²⁰. Estas arquitecturas, sin embargo, no han sido exploradas en volúmenes 3D+t, como las resonancias cardiacas, lo que implica una mayor complejidad debido a la reproducción tanto estructural como temporal de la secuencia.

En cuanto a GANS condicionales, en ²¹ se implementó la CB-GAN que toma como variable condicional mapas 2d del corte axial, con información espacial de zonas afectadas en el cerebro y permite generar formas emulando la patología condicional de entrada. También, se han reportado avances para generar artificialmente mo-

¹⁸ ARJOVSKY, Martin; CHINTALA, Soumith y BOTTOU, Léon. "Wasserstein gan". En: *arXiv preprint arXiv:1701.07875* (2017).

¹⁹ BERMUDEZ, Camilo, *et al.* "Learning implicit brain MRI manifolds with deep learning". En: *Medical Imaging 2018: Image Processing*. Vol. 10574. International Society for Optics y Photonics. 2018, pág. 105741L.

²⁰ GIHYUN KWON Chihye Han, Dae-shik Kim. "Generation of 3D Brain MRI Using Auto-Encoding Generative Adversarial Networks". En: *arXiv preprint arXiv:1908.02498* (2019).

²¹ MOK, Tony CW y CHUNG, Albert CS. "Learning data augmentation for brain tumor segmentation with coarse-to-fine generative adversarial networks". En: *International MICCAI Brainlesion Workshop*. Springer. 2018, págs. 70-80.

dalidades clínicas, a partir de otras secuencias de imágenes. Por ejemplo, en ^{22,23} se han podido generar tomografías, teniendo como referencia condicional las resonancias magnéticas. En ²⁴ se utilizó un esquema de parejas de entrenamiento, con tuplas (resonancia, tomografía) para generar resonancias magnéticas, proponiendo la arquitectura (MR-GAN), siendo efectiva para producir resonancias con coherencia estructural. En la figura 2 se ilustran algunos de los resultados obtenidos con algunas de estas arquitecturas (cycleGAN y MR-GAN) en el cambio de dominio, en este caso, para pasar de tomografías del cerebro a resonancias magnéticas y posteriormente reconstruir de nuevo la tomografía. Estas arquitecturas pueden ser importantes para la construcción de nuevas representaciones y análisis de biomarcadores en imágenes médicas.

En cuanto a cine-MRI cardíaco, formado por volúmenes temporales (3D+t), la generación de nuevos datos ha tenido mayores desafíos, relacionados con la dimensionalidad de los datos y las complejas estructuras cardíacas, acompañado por los movimientos no lineales registrados durante el ciclo cardíaco. Unas aproximaciones iniciales fueron propuestas bajo arquitecturas no-condicionales y un par generador-discriminador ²⁵. Alternativas más robustas han utilizado CycleGAN, arquitecturas que además usan diferentes modalidades para complementar la generación de ci-

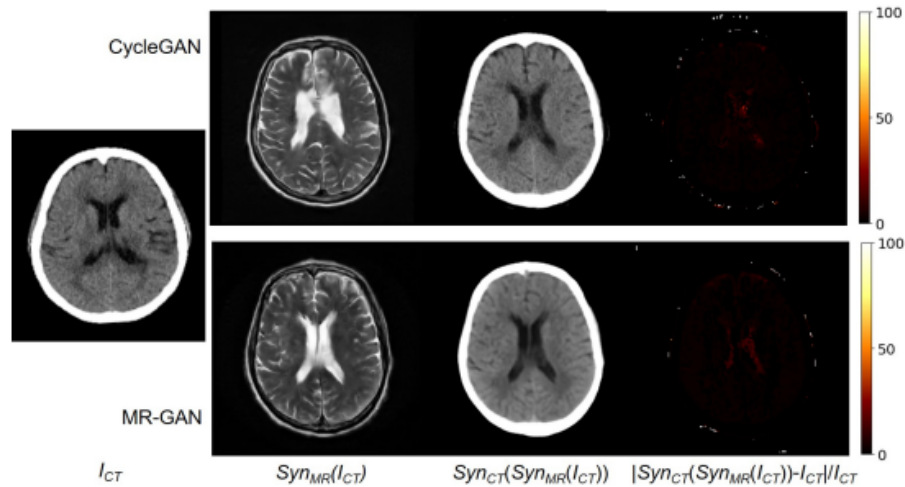
²² EMAMI, Hajar, *et al.* "Generating synthetic CTs from magnetic resonance images using generative adversarial networks". En: *Medical physics* 45.8 (2018), págs. 3627-3636.

²³ NIE, Dong, *et al.* "Medical image synthesis with context-aware generative adversarial networks". En: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2017, págs. 417-425.

²⁴ JIN, Cheng-Bin, *et al.* "Deep CT to MR synthesis using paired and unpaired data". En: *Sensors* 19.10 (2019), pág. 2361.

²⁵ ZHANG, Le; GOOYA, Ali y FRANGI, Alejandro F. "Semi-supervised assessment of incomplete LV coverage in cardiac MRI using generative adversarial nets". En: *International Workshop on Simulation and Synthesis in Medical Imaging*. Springer. 2017, págs. 61-68.

Figura 2. Cuadro comparativo CycleGAN vs MR-GAN



Fuente: Cheng-Bin JIN, *et al.* “Deep CT to MR synthesis using paired and unpaired data”. En: *Sensors* 19.10 (2019), pág. 2361

nes cardíacos²⁶. También, se han propuesto algunas aproximaciones para reconstruir volúmenes en 3D, intentando preservar la forma de las paredes cardíacas²⁷. Estos trabajos sin embargo, aún no tienen una suficiente fiabilidad para discriminar entre patologías o reproducir características particulares de ciertas anomalías cardíacas. En este sentido, esta área sigue siendo abierta y aún se trabajan en métodos para reproducir técnicas que modelen la complejidad del corazón durante el ciclo cardíaco.

²⁶ CHARTSIAS, Agisilaos, *et al.* “Adversarial image synthesis for unpaired multi-modal cardiac data”. En: *International Workshop on Simulation and Synthesis in Medical Imaging*. Springer. 2017, págs. 3-13.

²⁷ ZHANG, Zizhao; YANG, Lin y ZHENG, Yefeng. “Translating and segmenting multimodal medical volumes with cycle-and shape-consistency generative adversarial network”. En: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018, págs. 9242-9251.

1.2. BIOMARCADORES, VECTORES EMBEBIDOS Y REPRESENTACIONES

Un biomarcador digital es un indicador estructural o dinámico, que puede ser calculado utilizando diferentes sensores, y que permite la detección de enfermedades, correlación de patrones anormales, clasificación y valoración de los niveles de los estadios de una patología particular ²⁸.

En secuencias de imágenes médicas, los biomarcadores digitales comúnmente corresponden a patrones texturales, en términos espaciales o dinámicos (para secuencias temporales). Típicamente, esta información es modelada como descriptores, según la tarea a desarrollar o la información brindada por especialistas ²⁹. Recientemente, en el contexto de redes neuronales, se utilizan conjuntos masivos de datos para aprender representaciones visuales que puedan corresponder a una patología particular, o que solucionen una tarea específica. En cuanto a las redes convolucionales (discriminativas o generativas), los modelos logran representar la información visual en diferentes capas de abstracción, siendo los niveles superiores, vectores embebidos que recopilan de forma no-lineal y a través de variables ocultas la representación de cada una de las instancias y ejemplos en una entrada. Estos vectores embebidos son entonces representaciones de baja dimensionalidad de las variables de entrada que contienen características representativas de la variable en cuestión, las cuales aportan más información como entradas a modelos de clasificación en tareas supervisadas.

²⁸ ALTINTAS, Zeynep; FAKANYA, Wellington M. y TOTHILL, Ibtisam E. "Cardiovascular disease detection using bio-sensing techniques". En: *Talanta* 128 (2014), págs. 177 -186. DOI: <https://doi.org/10.1016/j.talanta.2014.04.060>.

²⁹ SAVADJIEV, Peter, *et al.* "Demystification of AI-driven medical image interpretation: past, present and future". En: *European Radiology* 29 (3 mar. de 2019), págs. 1616-1624. DOI: 10.1007/s00330-018-5674-x.

Los *Autoencoders* son arquitecturas dedicadas a la construcción de vectores embebidos para la representación de entradas altamente dimensionales ³⁰. Estos vectores embebidos han demostrado aprender relaciones típicas como lo obtenido por un análisis de componentes principales (PCA por sus siglas en inglés), pero además han logrado embeber relaciones no lineales que permiten una mejor representación de los objetos ³¹. Por ejemplo, en el trabajo de Jaime Zabalza³² se hace uso de *Autoencoders*, para la segmentación de una imagen hiper espectral en regiones a partir de sus vectores embebidos. Estas representaciones no sólo se pueden usar para clasificar, por ejemplo, Evgeny Putin presenta en un artículo de investigación³³, biomarcadores profundos del envejecimiento humano, en él, aplica la técnica de "stacking" la cual consta de entrenar múltiples modelos de clasificación, en este caso redes neuronales, con el fin de concatenar sus predicciones para obtener así un vector de representación, el cuál se usa para obtener mejores resultados.

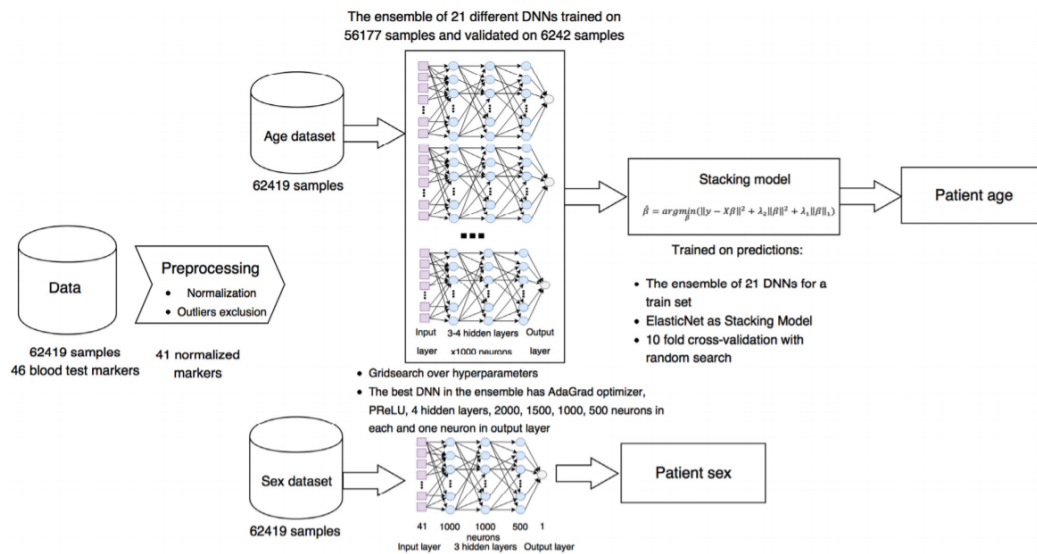
³⁰ KINGMA, Diederik P. y WELLING, Max. "An Introduction to Variational Autoencoders". En: *Foundations and Trends® in Machine Learning* 12.4 (2019), 307–392. DOI: 10.1561/22000000056.

³¹ MAKHZANI, Alireza, *et al.* *Adversarial Autoencoders*. 2016. arXiv: 1511.05644 [cs.LG].

³² ZABALZA, Jaime, *et al.* "Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging". En: *Neurocomputing* 185 (2016), págs. 1-10.

³³ PUTIN, Evgeny, *et al.* "Deep biomarkers of human aging: application of deep neural networks to biomarker development". En: *Aging (Albany NY)* 8.5 (2016), pág. 1021.

Figura 3. Pipeline del trabajo sobre desarrollo de biomarcadores por medio de redes neuronales



Fuente: Evgeny PUTIN, *et al.* “Deep biomarkers of human aging: application of deep neural networks to biomarker development”. En: *Aging (Albany NY)* 8.5 (2016), pág. 1021

Las arquitecturas convencionales también han usado vectores embebidos como *deep features* para representar objetos con métodos clásicos de aprendizaje de máquina, pero, también con la ventaja de adaptarlos a otro dominio sin el uso de entrenamientos exhaustivos ³⁴. Así mismo, en la literatura, se han reportado algunas aplicaciones que utilizan embebidos aprendidos en arquitecturas generativas y

³⁴ BORLAND, Lisa y HEEL, Marin van. “Classification of image data in conjugate representation spaces”. En: *J. Opt. Soc. Am. A* 7.4 (1990), págs. 601-610. DOI: 10.1364/JOSAA.7.000601.

han permitido ser usadas en otros dominios ³⁵³⁶³⁷.

1.3. DE PATRONES CARDIACOS A REPRESENTACIONES PROFUNDAS

Evaluar la función ventricular izquierda es una de las tareas que los cardiólogos suelen realizar a la hora de tomar decisiones en el manejo clínico de un paciente, análisis que se realiza con el objetivo de concluir si es necesario someter a un paciente a cirugías cardíacas, por esto es muy típico que se haga uso de medidas que facilitan el análisis de las patologías, siendo la fracción de eyección una medida la cual proporciona una aproximación confiable sobre el porcentaje de sangre que sale del corazón cada vez que se contrae por medio de la relación volumétrica del mismo en los momentos de sístole y diástole, ya que este parámetro está reconocido como una medida cuantitativa clínicamente útil del rendimiento ventricular, es importante su análisis a la hora de detectar patologías cardíacas; muchos de estos índices se pueden obtener a partir de la segmentación del ventrículo izquierdo de imágenes diagnósticas como lo son las resonancias magnéticas.

Es por esto que actualmente es común encontrar trabajos dedicados a aprender tareas de segmentación, como lo es con el trabajo de M.R. Avendi³⁸, el cual utili-

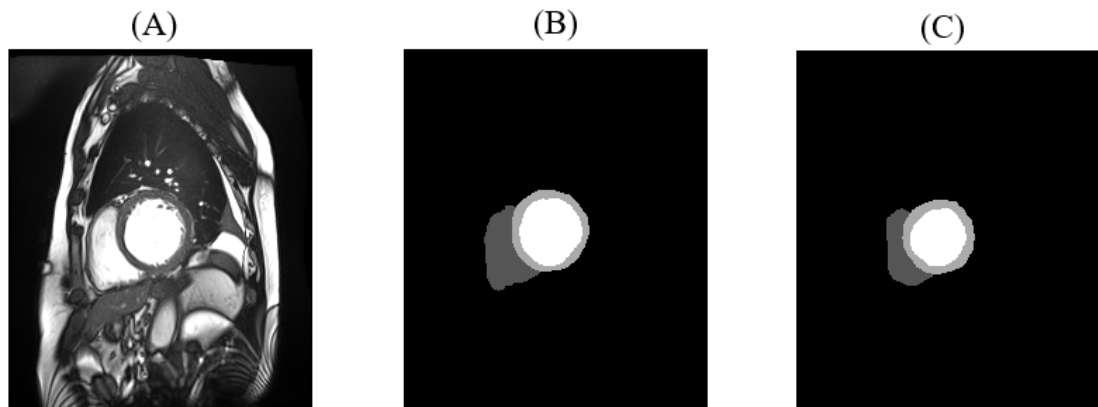
³⁵ MUKHERJEE, Sudipto, *et al.* "ClusterGAN: Latent Space Clustering in Generative Adversarial Networks". En: *Proceedings of the AAAI Conference on Artificial Intelligence* 33.01 (2019), págs. 4610-4617. DOI: 10.1609/aaai.v33i01.33014610.

³⁶ WU, Jiajun, *et al.* "Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling". En: *Advances in Neural Information Processing Systems*. Ed. por LEE, D., *et al.* Vol. 29. Curran Associates, Inc., 2016, págs. 82-90.

³⁷ VOYNOV, Andrey y BABENKO, Artem. *Unsupervised Discovery of Interpretable Directions in the GAN Latent Space*. 2020. arXiv: 2002.03754 [cs.LG].

³⁸ AVENDI, M.R.; KHERADVAR, Arash y JAFARKHANI, Hamid. "A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI". En: *Medical Image Analysis* 30 (2016), págs. 108 -119. DOI: <https://doi.org/10.1016/j.media.2016.01.005>.

Figura 4. (A) Resonancia magnética que ilustra las 2 cámaras ventriculares, (B) Segmentación del ventrículo izquierdo en el momento de la diástole isovolumétrica, (C) Segmentación del ventrículo izquierdo en el momento de la sístole isovolumétrica



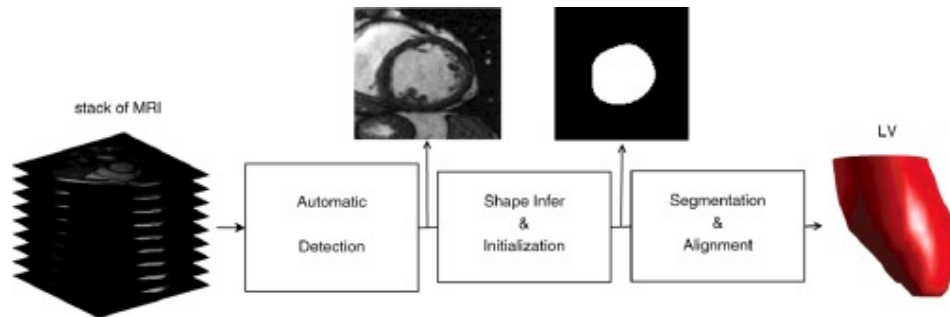
Fuente: Elaboración propia

za redes convolucionales para aprender la tarea de segmentación a partir de datos segmentados reales, luego valida la forma del ventrículo segmentado con *Autoencoders* para después utilizar modelos deformables con el fin de mejorar la precisión y la robustez de la segmentación. Por otro lado Ali Isin³⁹ en su trabajo, detecta arritmias cardiacas usando modelos de aprendizaje profundo en electrocardiogramas, con una metodología de Transfer learning con la arquitectura de Alexnet y 1,2 Millones de imágenes del conjunto de datos de ImageNet, esta metodología utiliza redes pre-entrenadas con el fin de reducir complejidad en el entrenamiento y aprovechar que estas saben estructuras locales o globales de la imagen.

En el dominio de Redes generativas adversarias existen muchos trabajos referentes

³⁹ ISIN, Ali y OZDALILI, Selen. "Cardiac arrhythmia detection using deep learning". En: *Procedia Computer Science* 120 (2017). 9th International Conference on Theory and Application of Soft Computing, Computing with Words and Perception, ICSCCW 2017, 22-23 August 2017, Budapest, Hungary, págs. 268 -275. DOI: <https://doi.org/10.1016/j.procs.2017.11.238>.

Figura 5. Esquema de red neuronal para segmentar



Fuente: M.R. AVENDI; Arash KHERADVAR y Hamid JAFARKHANI. “A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI”. En: *Medical Image Analysis* 30 (2016), págs. 108 -119. DOI: <https://doi.org/10.1016/j.media.2016.01.005>

al aumento de datos de imágenes diagnósticas debido a las limitaciones de estas, sumando que en el aprendizaje profundo se necesitan grandes cantidades de datos para obtener mejores resultados, aunque, no todos los trabajos apuntan al aumento de datos. En ⁴⁰ se propone una arquitectura GAN llamada Reconstruction Global-Local GAN (Recon-GLGAN) la cual reconstruye imágenes diagnósticas de tipo resonancia magnética. la arquitectura que se propone contiene una red U-NET y un discriminador de contexto que incorpora información global y local contextual de las imágenes que ofrecen un mejor desempeño en la tarea de reconstrucción. El método propuesto permite una mejor reconstrucción en general, condicionando esta con información previa de un área de interés.

⁴⁰ MURUGESAN, Balamurali, *et al.* “Recon-GLGAN: A Global-Local context based Generative Adversarial Network for MRI Reconstruction”. En: (ago. de 2019).

2. PLANTEAMIENTO Y JUSTIFICACIÓN DEL PROBLEMA

La representación de patrones cardiacos en secuencias cine-MRI es una tarea compleja debido a las estructuras no perimétricas del corazón, la limitada resolución de las imágenes y los diferentes patrones visuales que pueden variar entre una secuencia a otra.⁴¹ Sumado a lo anterior, el ciclo cardiaco es no lineal, con cambios bruscos que dificultan su modelamiento. Existen diferentes descriptores de corazón propuestos en la literatura, pero en general son enfocados a modelar características particulares de las paredes ventriculares. Con ello es posible medir patrones típicos usados en cardiología, pero limitando su exploración para obtener nuevas relaciones. Hoy en día, las arquitecturas generativas han permitido una representación robusta de espacio latentes, que permite calcular descriptores compactos que contienen relaciones relevantes y de alto nivel de los objetos modelados. Estos descriptores pueden guiar a la construcción de nuevos marcadores o al análisis de nuevas características distintivas en ciertas patologías cardiacas, lo cual podría apoyar tareas como el diagnóstico. Sin embargo, estos métodos hoy en día aprenden representaciones generales, no controladas, lo cual conlleva a generar imágenes, y en consecuencia vectores latentes, con poco sentido clínico. Por ejemplo, estos generadores pueden generar secuencias con información incoherente en cuanto a la morfología de los órganos que representa. Estas técnicas también pueden contribuir al aumento de datos, para que algoritmos de aprendizaje se ajusten de manera más fácil. Sin embargo la generación de datos incoherentes puede jugar en contra de esta hipótesis, conduciendo a modelos de aprendizaje incorrectos o limitados.

⁴¹ CHAPARRO-MUÑOZ, Marinela y PRASAD, Sanjay K. "Valor diagnóstico de la resonancia magnética cardiaca en la insuficiencia cardiaca: estado actual". En: *Cardiocre* 46.1 (2011), págs. 4-12.

3. OBJETIVOS

Objetivo general

Desarrollar un modelo generativo-adversario para la generación de resonancias magnéticas cardíacas condicionada por información relativa al corazón.

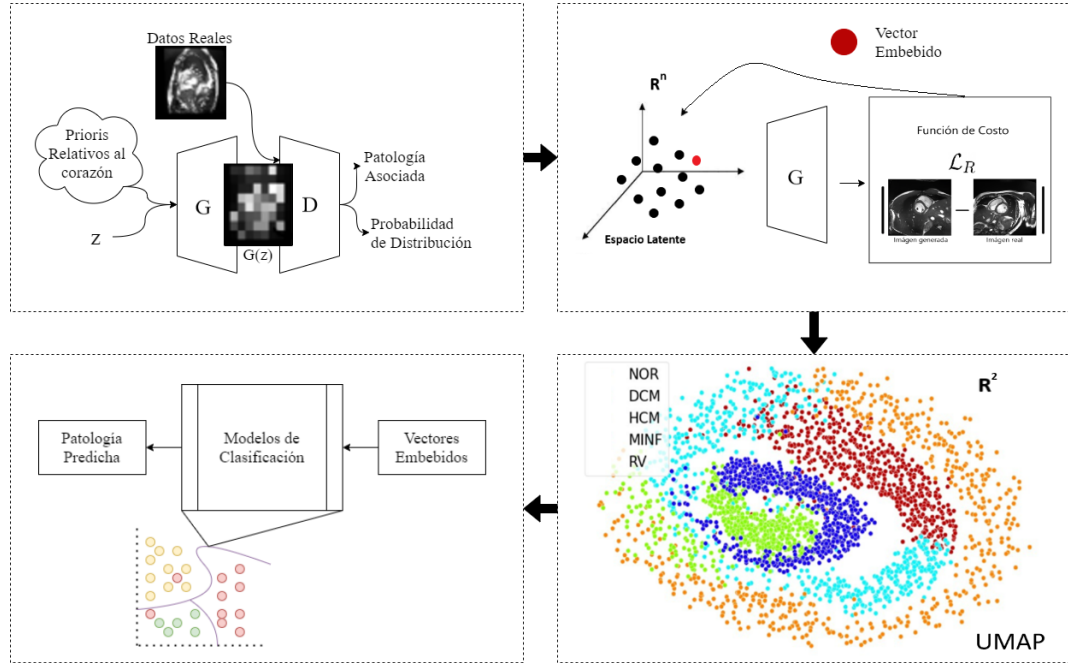
Objetivos específicos

- Seleccionar un conjunto de datos en el dominio de secuencias cardíacas para el entrenamiento del modelo y evaluación de hipótesis.
- Desarrollar un modelo generativo condicional en el dominio de la síntesis de secuencias de resonancias magnéticas cardíacas que integre los *prioris* establecidos.
- Implementar el modelo generativo establecido previamente.
- Evaluar los diferentes *a priori*s relacionados con la estructura y dinámica del corazón que puedan generar resultados morfológicamente válidos.
- Validar las secuencias producidas utilizando métricas de similitud aparente y relacionadas con conservación de la estructura de imágenes.

4. METODOLOGÍA PROPUESTA

En este trabajo se consideró el desarrollo, implementación y evaluación de una estrategia de aprendizaje profundo generativo, para la generación de secuencias cardíacas de cine-MRI, pero además, para el cálculo de vectores embebidos como descriptores de patologías particulares de entrenamiento. Como principal contribución del trabajo se propusieron diversas fuentes *a priori* de información cardíaca, con el fin de condicionar la representación obtenida. A partir de las imágenes cardíacas generadas se calcularon de forma iterativa vectores embebidos que permitieran resumir la información cardíaca en un esquema compacto de información. Estos vectores fueron agrupados en un esquema de baja dimensionalidad usando una estrategia de UMAP. Además, los puntos de baja dimensionalidad fueron evaluados cuantitativamente en cuanto a su capacidad de predecir patologías asociadas a la condición de entrada en el generador. El esquema general de la metodología propuesta es presentado en la figura 6. En las siguientes subsecciones se detallan los pasos propuestos en la metodología del presente trabajo.

Figura 6. Pipeline de la metodología propuesta.



Fuente: elaboración propia

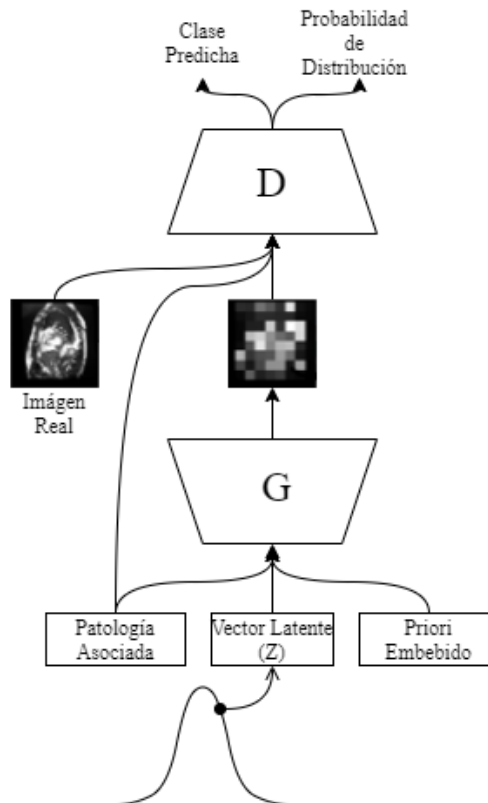
4.1. ESQUEMA GENERATIVO CON SOPORTE A MÚLTIPLES CONDICIONES

Las redes generativas adversarias han mostrado resultados impresionantes en el estado del arte para modelar y generar nueva información de un dominio particular⁴². Sin embargo, la versión general de estas arquitecturas tiende a presentar problemas por colapso de modo, es decir la generación de una única instancia o generar datos sobre una distribución específica, otro factor relevante, sobre todo en aplicaciones biomédicas, es la generación de artefactos o incoherencias estructurales sobre las secuencias generadas. Para evitar estas limitaciones, en la literatura se han propuesto diferentes esquemas generativos condicionados, los cuales permiten

⁴² SANTOS TANAKA, Fabio Henrique Kiyoyiti dos y ARANHA, Claus. *Data Augmentation Using GANs*. 2019. arXiv: 1904.09135 [cs.LG].

modelar problemas de múltiples instancias y controlar la generación de nuevas secuencias. En este trabajo se decidió implementar la red generativa MCIGAN, que entre otras permite modelar problemas de múltiples clases, con control en varias condiciones de entrada ⁴³. Este esquema generativo fue adaptado para la generación de imágenes de cine-MRI, como se ilustra en el esquema 7, con el fin de evaluar las distribuciones resultantes de diferentes *a priori* relativos al corazón.

Figura 7. Arquitectura de la red generativa adversaria condicional propuesta.



Fuente: elaboración propia

En términos generales, el método generativo MCIGAN opera con dos esquemas

⁴³ YILDIRIM, Gökhan; SEWARD, Calvin y BERGMANN, Urs. *Disentangling Multiple Conditional Inputs in GANs*. 2018. arXiv: 1806.07819 [cs.CV].

convolucionales típicos que actúan como generador y discriminador. Sin embargo, el generador está diseñado con tres entradas, las cuales representan una característica relativa a la distribución a generar. En este caso, estas entradas hacen referencia a un vector tridimensional normalizado entre $[-1, 1]$, que permite condicionar información de alto nivel de cada instancia, y regularizar la información generada con respecto a las etiquetas definidas en esta entrada. Por otra parte, existe una entrada que permite introducir un componente textural y es representado por un vector 512-dimensional definido por una distribución normal. Adicionalmente, esta arquitectura permite introducir de forma directa información estructural sobre regiones de interés en la imagen. En este sentido, se pueden introducir máscaras con la representación de la silueta, la cual se representa por medio de un vector 512-dimensional, producto de una red embebida cuyo input es una matriz de 64x64 píxeles. Para la información estructural se utilizó una red de segmentación que se describe en detalle en la siguiente sección. Esta red en particular se entrena de forma conjunta como un generador, permitiendo aprovechar una estrategia *end-to-end*. Estas tres entradas se conjugan en un único vector (1027-dimensional) que se pasan al generador para obtener $\tilde{x} = G(c, t, s)$, donde \tilde{x} es la imagen generada.

Una vez definido el esquema general de la arquitectura, se puede introducir la función de pérdida que permite ajustar los pesos de representación, tanto del generador, como del discriminador. En el caso de esta arquitectura la función es la distancia entre las distribuciones reales y generadas, también llamada función de costo de Wassertein, definida como:

$$\mathcal{L}_W = E_{x \sim P_r}[D'(x)] - E_{\tilde{x} \sim P_g}[D'(\tilde{x})] \quad (1)$$

donde L_W es la función de costo de Wassertein, $E_{x \sim P_r}$ es el valor esperado sobre todas las instancias de la distribución real de los datos r dado $D'(x)$ que es la es-

estimación del discriminador de la probabilidad de que la instancia de datos reales x sea real y $E_{\tilde{x} \sim P_g}$ el valor esperado sobre todas las imágenes generadas dado $D'(\tilde{x})$ que representa la estimación del discriminador de la probabilidad de que una instancia falsa sea real. En general, la función calcula la diferencia entre las probabilidades de imágenes en un dominio real vs falso.

El discriminador, además de evaluar si la imagen de entrada pertenece a un dominio real o falso, también estimará la condición c

$$A(x, s) = \frac{1}{|s|} \sum_{i,j} s(i, j) \cdot x(i, j) \quad (2)$$

donde $A(.,.)$ es la función que calcula el c promedio de la imagen (x), usando su respectiva máscara o segmentación (s), y $|s|$ el número de píxeles dentro de la segmentación.

4.1.1. CONSISTENCIA EN LAS CARACTERÍSTICAS Además de estas funciones, la arquitectura también cuenta con otras funciones de costos las cuales ayudan a la hora de hacer *tuning* por cada característica (c, t, s).

Por ejemplo, existe una función de costo para controlar la consistencia de la textura, para retener los detalles de una imagen de entrada por medio de matrices laplacianas, las cuales representan la estructura local alrededor de cada píxel y pueden ser utilizadas para medir la similitud estructural entre una imagen de origen y una de destino.

La consistencia de la forma se logra generando píxeles de fondo para las ubicaciones que están fuera de la máscara de segmentación.

En general se tiene

$$\min_G L_W + L_{aux} + \lambda_c L_c + \lambda_t L_t + \lambda_s L_s \quad (3)$$

en donde L representa cada función de costo mencionada anteriormente, y λ al peso que esta tendrá para el entrenamiento.

4.2. CONDICIONES CARDIACAS A PRIORI

Uno de los principales intereses en este trabajo es establecer las condiciones cardiacas que tienen una mayor correlación con ciertas patologías y que pueden ser modeladas dentro del esquema generativo para condicionar las imágenes generadas en la salida de la arquitectura. Teniendo en cuenta la información estructural y dinámica del corazón, en el presente trabajo se tuvieron en cuenta las siguientes fuentes de información:

4.2.1. SEGMENTACIÓN DEL CORAZÓN El ventrículo izquierdo (LV) en el corazón juega un rol fundamental en el sistema cardiovascular, siendo la localización de la mayoría de síntomas del daño cardiovascular. El estudio de su morfología y su comportamiento a través del ciclo cardiaco constituye una fuente de información primordial para la caracterización de las patologías cardiacas. En la rutina clínica dicho análisis inicia por la delineación de este ventrículo, por lo que es importante desarrollar técnicas de segmentación que faciliten su estudio, como se puede ver en la figura 4. La complejidad y variabilidad de la morfología cardiaca, el bajo contraste en la imagenología cardiaca, el ruido y la presencia de artefactos son algunas de las dificultades que caracterizan el proceso de segmentación del corazón como un problema abierto.

En la rutina clínica y el análisis radiológico a través de secuencias cine-MRI, una vez se tiene la segmentación del ventrículo en el fin de la diástole y el fin de la sístole se

pueden calcular ciertas características o indicadores patológicos como la masa miocárdica, grosor miocárdico y volumen ventricular ⁴⁴. Estas medidas son útiles para determinar la condición cardiaca del individuo y asociar posibles patologías según los hallazgos encontrados, lo que permite en cierta medida cuantificar y dar un diagnóstico imagenológico ⁴⁵. En la literatura, el problema de segmentación ventricular ha sido tratado desde múltiples enfoques que abarcan modelos estadísticos, modelos propagativos y más recientemente modelos basados en aprendizaje profundo ⁴⁶. Las técnicas basadas en *Autoencoders* reportan un mayor soporte y adaptabilidad a la variabilidad geométrica del ventrículo ⁴⁷. En este trabajo se decidió adoptar una técnica de aprendizaje profundo U-NET que permite aprender de forma general la segmentación más probable para un corte particular. La U-NET es una arquitectura de tipo *Autoencoder*, basada en la arquitectura para segmentación de imágenes médicas, presentada en la conferencia de la sociedad de computación de imágenes médicas e intervención asistida por computadora (MICCAI) del 2015 ⁴⁸. La cual consiste en dos redes, una contractiva, y una expansiva; la parte contractiva sigue una típica arquitectura convolucional, donde se van aumentando el número de canales

⁴⁴ ALFAKIH, Khaled, *et al.* "Assessment of ventricular function and mass by cardiac magnetic resonance imaging". En: *European radiology* 14.10 (2004), págs. 1813-1822.

⁴⁵ KANG, Dongwoo, *et al.* "Heart chambers and whole heart segmentation techniques: review". En: *Journal of Electronic Imaging* 21.1 (2012), págs. 1 -17. DOI: 10.1117/1.JEI.21.1.010901.

⁴⁶ VELÁZQUEZ, Edwin; BRAVO, Antonio y VERA, Miguel. "Segmentación del ventrículo izquierdo en imágenes cardíacas usando técnicas no paramétricas". En: mayo de 2015. DOI: 10.13140/RG.2.1.1910.2888.

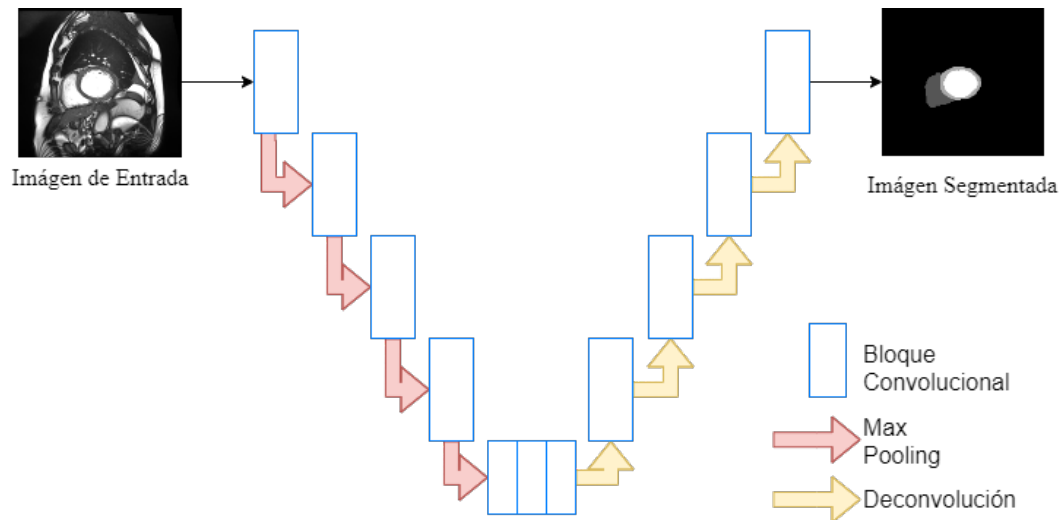
⁴⁷ BIFFI, C., *et al.* "3D High-Resolution Cardiac Segmentation Reconstruction From 2D Views Using Conditional Variational Autoencoders". En: *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. 2019, págs. 1643-1646. DOI: 10.1109/ISBI.2019.8759328.

⁴⁸ RONNEBERGER, Olaf; FISCHER, Philipp y BROX, Thomas. "U-Net: Convolutional Networks for Biomedical Image Segmentation". En: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. por NAVAB, Nassir, *et al.* Cham: Springer International Publishing, 2015, págs. 234-241.

de características con el fin de aprender una representación, de lo que en este caso sería la segmentación. Por otra parte, en la parte expansiva se hará lo contrario; en ella se buscará reducir a la mitad el número de canales de características mientras se aumenta su dimensión con canales de convolución ascendente, con el fin de permitir una localización precisa de la información aprendida en la parte contractiva. En la capa final se usa una convolución para mapear cada vector de característica al número deseado de la imagen de salida.

Esta arquitectura fue entrenada con segmentaciones calculadas únicamente en dos momentos del ciclo cardiaco (fin del diástole y fin del sístole). En la figura 8 se ilustra la arquitectura U-NET adaptada en este trabajo para codificar la información morfológica del ventrículo izquierdo, con el fin de ser usados como *priori* para la generación de resonancias magnéticas sintéticas.

Figura 8. Arquitectura U-NET usada para la segmentación de resonancias magnéticas cardiacas.



Fuente: elaboración propia

Se utilizó el *Dice-Score* como medida de validación para saber si la tarea de seg-

mentación fue correcta, básicamente, es una medida de superposición entre dos imágenes, o en este caso, segmentaciones. Esta medida va de 0 a 1, donde un coeficiente de 1 denota un solapamiento perfecto y completo. La función es la siguiente:

$$f_{Dice} = \frac{2|A \cap B|}{|A| + |B|} \quad (4)$$

Donde $|A \cap B|$ representa los elementos comunes entre los conjuntos A y B , y $|A|$ representa el número de elementos en el conjunto A (lo mismo sería para el conjunto B).

Figura 9. Ejemplo de una segmentación aprendida por el modelo U-NET.



Fuente: elaboración propia

4.2.2. FLUJO ÓPTICO Además de la configuración morfológica de las paredes cardiacas, la morfofisiología del corazón es un biomarcador fundamental para la

caracterización de condiciones anormales o patológicas⁴⁹. Esta información esta directamente relacionada con la deformación temporal de los ventrículos y su capacidad sincronizada para bombear ciertos volúmenes de sangre. El flujo óptico es la primera alternativa para cuantificar patrones de movimiento en secuencias de vídeo. En esencia, el flujo óptico mide el desplazamiento local de los píxeles entre dos cuadros consecutivos, resultando en un mapa vectorial con información relacionada con la velocidad aparente de los objetos en movimiento. En este trabajo se decidió cuantificar el movimiento del corazón como una medida de movimiento aparente calculada a través de un algoritmo de flujo óptico⁵⁰.

En la literatura existe una amplia variedad de algoritmos de flujo óptico para diferentes aplicaciones y utilizando diferentes restricciones para asociar los píxeles entre cuadros consecutivos. Sin embargo, la mayoría de alternativas tienen restricciones fuertes que solo consideran coherencia de movimiento en desplazamientos pequeños entre cuadros consecutivos. Esta suposición puede ser restrictiva para la cuantificación de movimientos del corazón, considerando que durante el ciclo cardiaco existen movimiento abruptos y condiciones no lineales que hacen parte de movimientos típicos. Para este trabajo se hizo uso de un flujo óptico denso de largos desplazamientos de movimiento⁵¹, a escala de píxel, con el fin de detectar los movimientos salientes asociados a fuertes desplazamientos entre cortes consecutivos t y $t + 1$, lo cual lleva a una representación de la velocidad de las cavidades cardiacas

⁴⁹ GOKSEL, D.; OZKAN, M. y OZTURK, C. "Cardiac motion analysis in MRI for classification". En: *Proceedings IEEE International Symposium on Biomedical Imaging*. 2002, págs. 935-938. DOI: 10.1109/ISBI.2002.1029415.

⁵⁰ MEDINA, Ruben, *et al.* "Cardiac Motion Estimation in Magnetic Resonance Images Using Optical Flow". En: *IEEE Latin America Transactions* 14 (jun. de 2016), págs. 2807-2816. DOI: 10.1109/TLA.2016.7555259.

⁵¹ BROX, T.; BREGLER, C. y MALIK, J. "Large displacement optical flow". En: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 2009, págs. 41-48. DOI: 10.1109/CVPR.2009.5206697.

en este par de cortes temporales.

Este flujo óptico considera restricciones típicas de densidad entre cuadros consecutivos expresados por I_t y I_{t+1} , para poder localizar el píxel temporalmente. Estas restricciones, básicamente, asocian un píxel temporalmente, según la información estructural y de intensidad, expresada como los bordes. Las medidas de restricción utilizadas pueden definirse como:

$$E_c(v) = |I_t(x + v(x)) - I_{t+1}(x)|^2 \quad (5)$$

$$E_g(v) = |\nabla I_t(x + v(x)) - \nabla I_{t+1}(x)|^2 \quad (6)$$

donde v es el campo de movimiento resultante, dado por el vector de desplazamiento $(u, v)^T$, para cada píxel x , con cortes en los momentos temporales, siendo E_c la restricción típica de color o intensidad y E_g la restricción de gradiente invariante a los cambios de brillo ($\frac{dE}{dt} = 0$).

Como criterio adicional, este trabajo también introduce una restricción para considerar la información de movimiento aparente en cada localización. Esta información es introducida como una nueva regla de minimización, expresada como:

$$E_s(v) = |\nabla u(x)|^2 + |\nabla v(x)|^2 \quad (7)$$

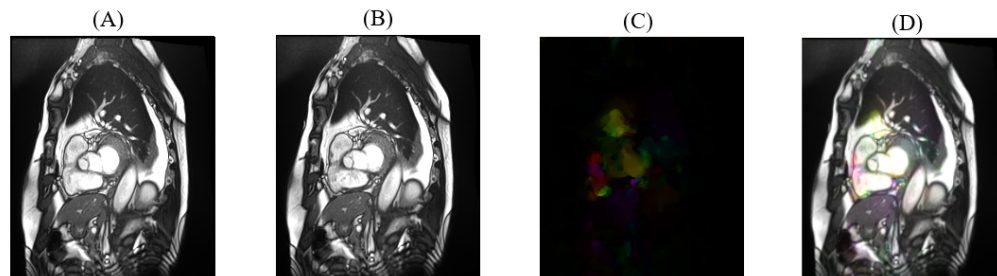
Finalmente, la estrategia aquí propuesta, implementa una restricción no local, que le permite asociar movimientos largos, entre cuadros consecutivos, asociando píxeles similares temporalmente. Esta restricción utiliza operadores no locales como los descriptores SIFT ($f_t(x)$) para asociar píxeles no-locales similares y asociar un

vector de movimiento aparente. Esta nueva regla de restricción esta definida como:

$$E_d(v_1) = |f_{t+1}(x + v_1(x)) - f_t(x)|^2 \quad (8)$$

donde d representa el vector descriptor y f los patrones de velocidad de las regiones no locales en los tiempos $(t, t + 1)$.

Figura 10. (A) Resonancia magnética en un momento t , (B) Resonancia magnética en un momento $t + 1$, (C) Flujo óptico resultante de los momentos $(t, t + 1)$, (D) Flujo óptico sobre la resonancia magnética en la instancia t

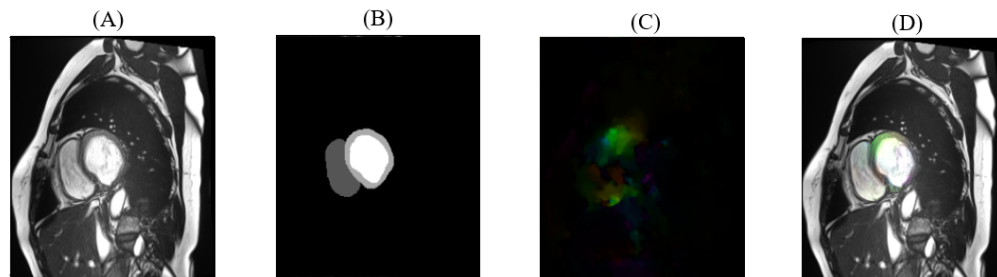


Fuente: elaboración propia.

Una limitación de la implementación cruda del flujo óptico, sobre todo en la imagen de la resonancia magnética, es la consecuente colección de movimientos ruidosos y no asociados al patrón cardiaco. Estadísticamente, estos segmentos de ruido, de baja intensidad, pero obtenidos en regiones espacialmente significativas de la imagen pueden limitar el aprendizaje de los patrones del corazón referentes a una patología. Para mitigar este problema, en este trabajo se utilizó la segmentación del corazón como máscara morfológica, siendo el producto con la segmentación, los patrones de movimiento únicamente asociado a los ventrículos. Este flujo focalizado se utilizó como *priori* en la generación de resonancias magnéticas cardiacas sintéticas. En la figura 11 se ilustran los flujos ópticos resultantes para una secuencia típica, tratada

en este trabajo.

Figura 11. (A) Resonancia magnética en un momento t , (B) Corazón segmentado de la resonancia magnética en la instancia t , (C) Flujo óptico resultante de los momentos $(t, t+1)$, (D) Flujo óptico segmentado sobre la resonancia magnética en la instancia t



Fuente: elaboración propia.

4.2.3. PATOLOGÍA ASOCIADA En el trabajo también se consideró la información brindada por expertos cardiólogos y anotadas como patologías generales que definían la condición de una secuencia de cine-MRI. Entonces, otro *a priori* con el que se cuenta para la generación de secuencias cardiacas son las patologías, las cuales sirven como etiqueta para diferenciar las variaciones en la estructura del corazón. A continuación, se mencionará con más detalle cada una de estas patologías y los parámetros fisiológicos más importantes al momento de diferenciarlas.

Para la clasificación y etiquetado de estas condiciones se tuvo en cuenta medidas fisiológicas de capacidad de bombeo del corazón. Uno de estos parámetros es la fracción de eyección, el cual mide el porcentaje de sangre que sale del corazón cada vez que se contrae, este suele medirse con el ventrículo izquierdo el cual se puede definir como la cámara de bombeo principal del corazón. Este valor se puede

calcular con la siguiente fórmula

$$\mathcal{FE} = \frac{VDF - VSF}{VDF} * 100 \quad (9)$$

donde VDF hace referencia al volumen diastólico final, que mide el volumen de sangre que llena el ventrículo del corazón al final de la fase de relajación del mismo; y donde VSF se refiere al volumen sistólico final, el cual mide el volumen de sangre en el ventrículo cuando el corazón se contrae. Los valores normales están entre 55 y 100 %.

Las patologías consideradas para este proyecto son:

- Infarto de miocardio previo cuyas características son: fracción de eyección del ventrículo izquierdo inferior al 40 % y varios segmentos de miocardio con contracción anormal (**MINF**).
- Cardiomiopatía dilatada con volúmenes ventriculares izquierdos diastólicos $> 100\text{mL}/\text{m}^2$ y una fracción de eyección del ventrículo izquierdo inferior al 40 % (**DCM**).
- Miocardiopatía hipertrófica que se caracteriza por tener masa cardiaca del ventrículo izquierdo superior a $110\text{g}/\text{m}^2$, varios segmentos de miocardio con un grosor superior a 15mm en diástole y una fracción de eyección normal (**HCM**).
- Ventrículo derecho anormal con volúmenes de la cavidad ventricular derecha superior a $110\text{mL}/\text{m}^2$ o fracción de eyección del ventrículo derecho inferior al 40 % (**RV**).

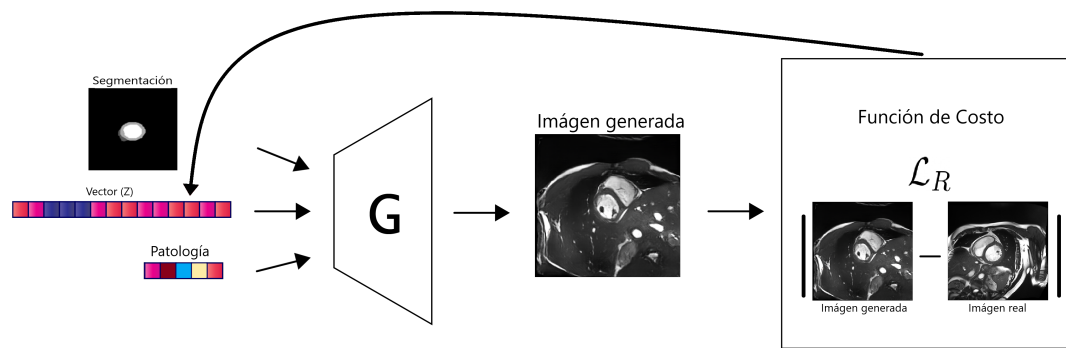
Además de estas patologías también se tienen en cuenta pacientes con corazón normal (**NOR**).

4.2.4. GENERACIÓN DE EMBEBIDOS Uno de los principales intereses de este trabajo es la descripción compacta de patrones cardiacos asociados a ciertas patologías de entrenamiento o condiciones *a priori* particulares. También se busca que la representación profunda pueda brindar nuevos patrones, codificados en la representación jerárquica y compactados en los últimos niveles de representación. En estos niveles (capas de la red) se pueden tener entonces vectores embebidos, cuyas activaciones pueden representar nuevos biomarcadores cardiacos, los cuales pueden tener una asociación directa con las patologías desarrolladas.

Los vectores embebidos en las redes profundas convencionales suelen obtenerse de forma fácil después de operar hacia adelante una imagen convencional. Sin embargo, en los esquemas generativos, diseñados en generar imágenes, los vectores embebidos resultantes deben ser operados en la primera capa de la red, a partir de la imagen generada, es decir, haciendo una retro-propagación de los valores obtenidos, hacia el vector de entrada. Una vez entrenada la arquitectura generativa adversaria, entonces se procede a utilizar los datos de validación como imágenes objetivo para la búsqueda de los vectores embebidos z que más se acerquen a una imagen $G(z)$ de la distribución aprendida por el generador. Esta aproximación se hace a partir de un vector aleatorio, donde después de manera iterativa, por medio de una función de costo como fin de la iteración, se va ajustando el vector para obtener la imagen objetivo, como se puede evidenciar en la figura 12. Esta función de costo está constituida por una diferencia absoluta matricial.

4.2.5. AGRUPAMIENTO DESDE UMAP Después de obtener los vectores embebidos de los datos de validación, se puede representar cada una de las patologías según la descripción de estos vectores. Una estrategia para comprobar su capacidad de representar información cardiaca coherente es mapearla a un espacio de baja dimensionalidad y evaluar la capacidad de agrupar vectores con similar diag-

Figura 12. Obtención de vector en el espacio latente ajustado por una resta absoluta entre la imagen objetivo y la generada $||r - l_g||$



Fuente: elaboración propia.

nóstico. Además, esta estrategia permite de forma directa introducir una herramienta de soporte al diagnóstico que permite visualizar poblaciones de muestras con ciertas características patológicas y al mismo tiempo mapear nuevos vectores, para que un experto cardiólogo pueda ver asociaciones con respecto a ciertas poblaciones características. De esta forma, una visualización de estos vectores embebidos puede ser altamente relevante como estrategia de soporte en la valoración, diagnóstico y seguimiento de pacientes.

Para la reducción dimensional, proyección y agrupamiento, en este trabajo se decidió implementar la estrategia de UMAP (por sus siglas en inglés *Uniform Manifold Approximation and Projection*) (Aproximación y proyección de manifold uniformes)⁵². UMAP es un algoritmo de reducción de dimensión de propósito general creado en 2018, el cual cuenta con múltiples hiper-parámetros tanto para determinar la dimensionalidad del espacio de dimensión reducida de salida como para controlar cómo el algoritmo trata la estructura local y global de los datos. En esta estrategia se considera que los datos están uniformemente distribuidos, que la métrica en

⁵² MCINNES, Leland; HEALY, John y MELVILLE, James. *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction*. 2020. arXiv: 1802.03426 [stat.ML].

alguna geodésica es constante para distancias pequeñas, y el *manifold* es localmente conectado. Entonces, los vectores en alta dimensión (vectores embebidos) son proyectados a un espacio de *Riemman* (espacio curvo, construido típicamente a partir de correlaciones entre características usando matrices simétricas positivas). En este espacio topológico de *Riemman*, los puntos son conectados según las geodésicas convenientes para este espacio y su correspondiente asociación es definida nuevamente en el espacio original. Cabe acotar, que para la metodología propuesta de este trabajo, se podría utilizar cualquier técnica de reducción dimensional, como por ejemplo, los análisis de componentes principales o t-sne. Sin embargo, una experimentación ágil, sobre algunos experimentos iniciales, en el presente trabajo, evidenciaron una sobresaliente descripción de los vectores embebidos, utilizando la estrategia de UMAP.

4.2.6. ESTRATEGIAS DE CLASIFICACIÓN Una vez los vectores embebidos son proyectados en un espacio topológico, las características de correlación resultantes pueden mapearse a diferentes clasificadores para validar cuantitativamente la capacidad de predecir patologías. En este trabajo entonces se utilizaron estrategias típicas de aprendizaje de máquina para validar la capacidad de esta representación topológica y cada uno de sus puntos correspondientes para representar una patología particular. A continuación, se presentará un resumen general de las técnicas utilizadas y su principal hipótesis para la definición de las fronteras de clasificación.

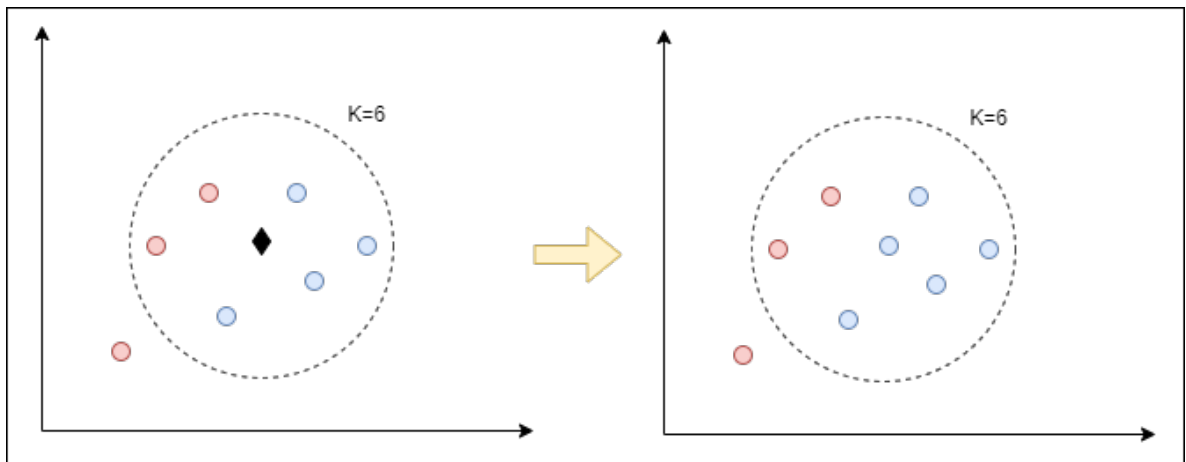
4.2.7. K-VECINOS CERCANOS Uno de los modelos usados fue el K-Nearest Neighbors(KNN), el cual es un modelo de clasificación supervisado, es decir, que usa la característica relativa a su clase asociada, en este caso los embebidos y su patología, para hacer predicciones con futuros nuevos datos de entrada. El modelo en general clasifica a partir de los datos más cercanos a un punto; esta distancia se

calcula de 3 maneras, pero la más general es la distancia euclidiana

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^p (x_{ri} - x_{rj})^2} \quad (10)$$

dada por la resta "ordinaria" entre dos puntos de un espacio euclídeo p , en este caso representados por x_i y x_j . El modelo empieza con un punto aleatorio del espacio y calcula los k puntos más cercanos, clasificando a este en la categoría cuya mayor cantidad de puntos tenga alrededor, tal como se ve en la figura 13. A pesar de su simplicidad, una de las principales fortalezas de este método, es la construcción de fronteras no paramétricas, lo cual puede ser crucial cuando se desconoce la distribución que siguen ciertos puntos para una clase particular.

Figura 13. Representación del proceso de clasificación de un nuevo punto con el modelo de KNN con k igual a 6.



Fuente: elaboración propia.

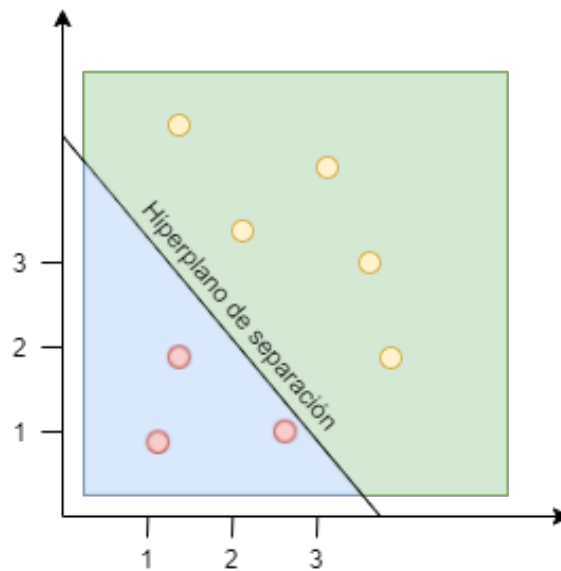
4.2.8. MÁQUINAS DE SOPORTE VECTORIAL Las Máquinas de Soporte vectorial (SVM, *Support vector machine*, por sus siglas en inglés) son otro conjunto de algoritmos comúnmente usados para tareas de clasificación. Estos funcionan por medio de los denominados hiperplanos, los cuales son subespacios $R - 1$ dimen-

sionales, siendo R el espacio en el cual se encuentran los datos. Por ejemplo, en el caso bidimensional, estos subespacios se podrían describir acorde a la ecuación de una recta

$$\lambda_0 + \lambda_1 x_1 + \lambda_2 x_2 = 0 \quad (11)$$

en donde los parámetros λ son los que componen el hiperplano, el cual, en la tarea de clasificación, serviría para separar el espacio en distintos subespacios. En la figura 14 puede verse un ejemplo de separación del espacio por medio de un hiperplano lineal.

Figura 14. Ejemplo de hiperplano de separación en un espacio bidimensional.



Fuente: elaboración propia.

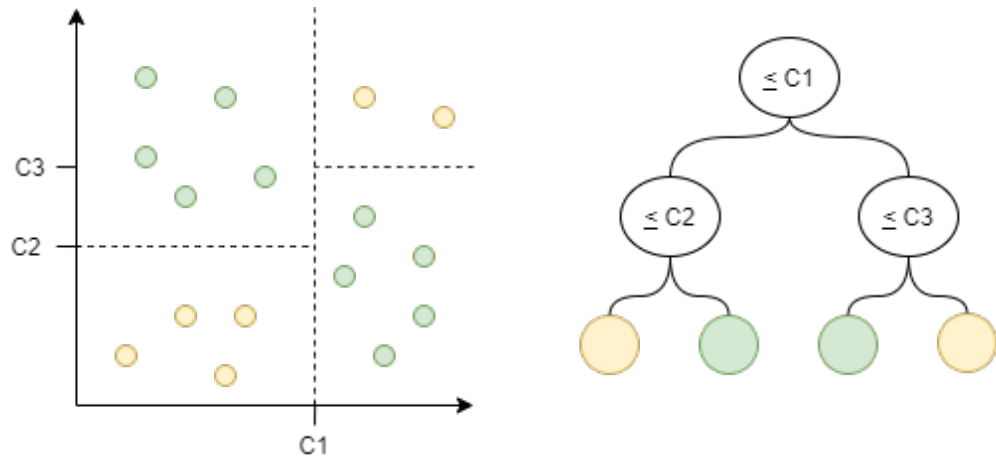
Una de las principales ventajas de los SVM es que pueden definir nuevas relaciones y dimensiones, cuando los datos no se pueden separar en un hiperplano original. En este caso se hace uso de una transformación que permite aumentar las dimensiones, utilizando kernels, que son funciones que actúan de forma local, entre las características originales para producir nuevas dimensiones. Estas funciones o kernels del SVM, son el resultado del producto punto de dos vectores en un nuevo espacio dimensional distinto al original en el que se encuentran los vectores. Entre los kernels más utilizados se encuentran el lineal, el polinómico, y el RBF (funciones de base radial). Esta estrategia ha permitido categorizar a los SVM como el estado del arte en muchos de los problemas de aprendizaje de máquina, pero la búsqueda de sus hiper-parámetros, selección de kernels, y ajuste de funciones puede ser una tarea tediosa y sus funciones de frontera pueden ser sensibles en cada conjunto de datos.

4.2.9. BOSQUE ALEATORIO Teniendo en cuenta la complejidad de los vectores embebidos que representan cada cine-MRI, en este trabajo se tuvieron en cuenta diferentes estrategias, con diferentes hipótesis de partición. Dentro de las estrategias implementadas, también se validó la partición discreta utilizando árboles de representación entre características. Los bosques aleatorios (RF) son hoy en día la estrategia que permite mitigar la sensibilidad de los árboles de decisión, como representación única. Para entender los bosques aleatorios (RF) primero se debe abarcar lo que es un árbol de decisión.

Los árboles de decisión son algoritmos de clasificación supervisado en donde se divide un espacio en diferentes regiones a partir de una serie de condiciones aprendidas a partir del conjunto de características de entrenamiento. En otras palabras, el espacio se divide en conjuntos homogéneos, a partir de las variables de entrada más significativas. Estas condiciones siguen una ramificación, la cual le da nombre

al modelo, con el fin de optimizar el proceso de división analizando las variables en un orden de relevancia; tal como se ve en la figura 15.

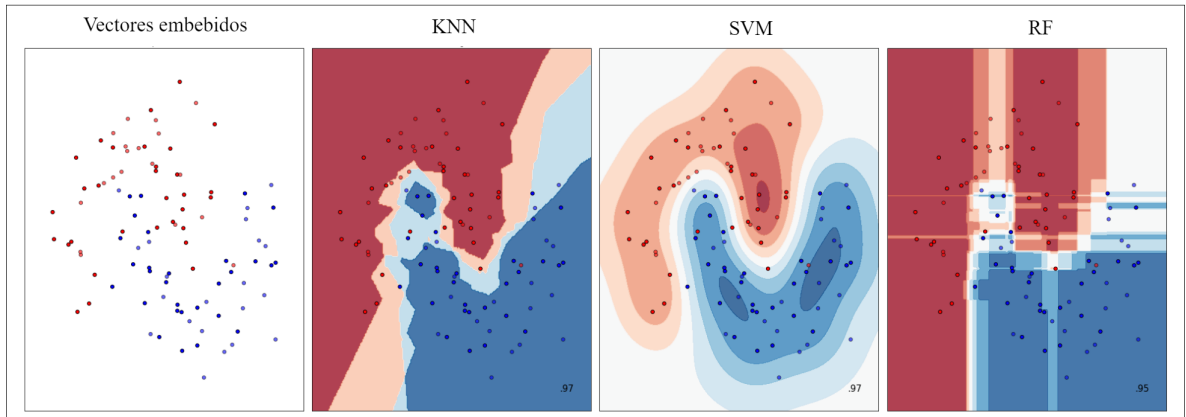
Figura 15. Ejemplo de clasificación por medio de un árbol de decisión junto con su diagrama.



Fuente: elaboración propia.

Entonces, los bosques aleatorios no son más que un conjunto de árboles de decisión, combinados en un modelo robusto; en donde cada árbol da una clasificación, siendo la predicción resultante la clase con mayor número de votos en el conjunto de árboles.

Figura 16. Ejemplo de clasificación para distintos modelos de aprendizaje.



Fuente: elaboración propia.

En la figura 16 se ilustran algunos ejemplos de la partición resultante en el espacio de embebidos utilizando las diferentes estrategias de clasificación mencionadas anteriormente, las cuales dan como resultados diferentes separaciones de acuerdo a las hipótesis establecidas para las fronteras.

4.3. DISEÑO EXPERIMENTAL

4.3.1. SELECCIÓN DE DATOS Para el desarrollo del proyecto fue usado el dataset público del desafío (The MICCAI 2017 automated cardiac diagnosis challenge (ACDC)) el cual consta de 150 pacientes donde cada uno cuenta con volúmenes Cine-MRI con vistas de eje corto; se utilizaron únicamente 100 pacientes de los totales ya que estos son los únicos que contienen una segmentación y clasificación de a cuál grupo pertenecen.

También, se utilizó sólo el 70% de los *slíces* por pacientes, con el fin de descartar los *slíces* apicales, ya que estos aportan una caracterización menos confiable a su patología asociada, por lo que al final se contó con un total de 16813 instancias para

el entrenamiento de la red generativa.

Los datos fueron divididos en 2 conjuntos distintos: 80 % para entrenamiento de la MCIGAN, y 20 % para validación y generación de espacios latentes. los conjuntos de datos fueron distribuidos de tal manera que se evitara un desbalance de clase y un colapso de modo en la arquitectura GAN.

4.3.2. CONFIGURACIÓN DE LA MCIGAN Se establecieron una dimensión de 64x64 tanto para la salida como para la entrada de la red generativa, las entradas del generador son 3, un vector de espacio latente de dimensión 512, una imagen de tamaño 64x64 referente a la segmentación, la cual se codifica en el entrenamiento en un vector 512 dimensional, y una entrada 3-dimensional que hace referencia a la condición *c*, que en este caso es un mapeo de la patología asociada.

se hicieron en total 3 pruebas, con el fin de comparar los resultados o el aporte de cada uno de los *prioris* al momento del entrenamiento, en la primera se utilizó la segmentación, vector en el espacio latente, y clase a la cual pertenece, mientras que en el segundo experimento se cambió la segmentación por el flujo, y en el tercero se utilizó el flujo en el área de la segmentación junto con la patología asociada.

Para el entrenamiento de la GAN, se hizo uso de 1 GPU **Nvidia Titan V**, con *minibatches* de 32 imágenes de entrada, la duración del entrenamiento se mide de la misma forma que la arquitectura ProgressiveGAN⁵³, siendo esta en miles de imágenes, para estas 3 pruebas el valor fue de 10000, lo cual la GPU usada se demoró en procesar aproximadamente 5 días por cada experimento. los demás hiperparáme-

⁵³ KARRAS, Tero, *et al.* "Progressive growing of gans for improved quality, stability, and variation". En: *arXiv preprint arXiv:1710.10196* (2017).

tros se dejaron igual al del artículo original. Además, para el entrenamiento de los 3 modelos se usó un peso de 100 en las consistencias de la GAN; A pesar de usar una arquitectura progresiva, debido a la inclusión de múltiples entradas, y modificación de las funciones de costo, no se utilizó un enfoque progresivo.

4.3.3. VECTORES EN ESPACIO LATENTE Los vectores se calcularon en *batches* de 64 para optimizar el proceso de cálculo, el proceso de ajuste se hizo de manera iterativa y los criterios de parada fueron un *loss value* menor a 0.05 o 5000 iteraciones.

4.3.4. VALIDACIÓN ESTADÍSTICA Al momento de reducir la dimensionalidad por medio del algoritmo de UMAP, se utilizó métrica binaria de *hamming*, con una salida de 2 componentes, y al momento de validar estos vectores embebidos resultantes de la reducción de dimensión, se utilizaron 3 clasificadores, los cuales fueron ajustados con una optimización bayesiana para obtener los hiper parámetros que mejor se ajustaran a los datos de entrada.

A continuación, se mostrarán los mejores hiper-parámetros de los modelos, obtenidos a partir de una estrategia de búsqueda exhaustiva con el fin de obtener los mejores resultados de clasificación; esta búsqueda exhaustiva se hizo por medio de un *Tuner*, el cual actúa de manera iterativa, buscando en un rango o conjunto de parámetros dados, los hiper-parámetros que mejor se ajusten a cada modelo.

Cuadro 1. Hiperparámetros de entrenamiento para el modelo de Knn

| Condiciones | Tamaño Hojas | Número de Vecinos | Valor de P |
|-----------------------------|---------------------|--------------------------|-------------------|
| Segmentación + condición(c) | 1 | 15 | 2 |
| Flujo + c | 1 | 22 | 2 |
| Segmentación + Flujo + c | 1 | 13 | 2 |

Cuadro 2. Hiperparámetros de entrenamiento para el modelo de Random Forest

| Condiciones | Número de estimadores |
|--------------------------|------------------------------|
| Segmentación + c | 54 |
| Flujo + c | 74 |
| Segmentación + Flujo + c | 115 |

Cuadro 3. Hiperparámetros de entrenamiento para el modelo de support vector machine.

| Condiciones | Kernel | Grados |
|--------------------------|---------------|---------------|
| Segmentación + c | rbf | 3 |
| Flujo + c | rbf | 4 |
| Segmentación + Flujo + c | rbf | 4 |

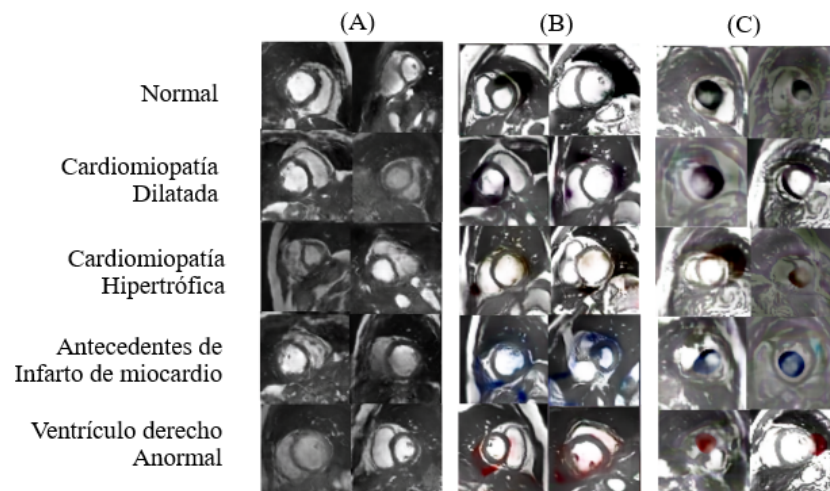
5. EVALUACIÓN Y RESULTADOS

La metodología aquí propuesta fue validada con un conjunto de datos públicos, categorizada con diferentes condiciones cardiacas, utilizando una representación generativa condicionada por diferentes fuentes de información. Para la validación de la metodología propuesta se configura la arquitectura MCIGAN, se ajustó su representación para modelar secuencias de cine-MRI, y se validaron diferentes restricciones para evaluar la capacidad de codificar descriptores de condiciones cardiacas. Una vez entrenada la arquitectura convolucional, se calcularon vectores embebidos, se proyectaron en un espacio de baja dimensionalidad y se procedió a evaluar su capacidad de representación, dentro de un marco de clasificación supervisada.

Los resultados cualitativos relacionados con la capacidad de generar imágenes se pueden ilustrar en la Figura 17. En esta figura se pueden observar imágenes obtenidas para diferentes patologías y condicionadas con diferente información morfológica, obtenida con el flujo óptico e integrando tanto información relacionada con la segmentación como con el movimiento cardiaco. En términos observacionales, cabe resaltar la calidad significativa lograda por la arquitectura cuando su única condición esta relacionada con la estructura de los ventrículos (columna a la izquierda). Este resultado puede estar asociado a una restricción fuerte en la generación de ventrículos con coherencia espacial, siendo una distribución de buena calidad, pero homogénea entre las diferentes condiciones cardiacas. En la columna central se pueden observar algunos ejemplos relacionados a imágenes generadas utilizando la información dinámica del corazón. En este caso, las regiones en color representan centros de atención con fuerte correlación a la patología de entrenamiento. Si bien estas imágenes tienen algunas inconsistencias espaciales, como se verán en los resultados cuantitativos, resultan ser interesantes para describir patrones diferenciales entre las patologías. Del mismo modo, una vez se integran las dos informaciones a

priori (espacial y temporal) la generación de imágenes no tiene condiciones óptimas, pero su vector embebido puede cuantificar y enfocarse en patrones que permiten un estudio de patrones ocultos relacionados con condiciones cardiacas, y por ende son las de mayor interés en nuestro estudio.

Figura 17. (A) imágenes generadas con entradas de segmentación y patología asociada, (B) imágenes generadas con entradas de flujo óptico y patología asociada, (C) imágenes generadas con entradas referentes al flujo en región segmentada junto con su patología.



Fuente: elaboración propia.

En las siguientes subsecciones se presenta un análisis cuantitativo de las representaciones embebidas obtenidas con la red generativa entrenada utilizando y validando diferentes fuentes de información cardiaca. Como se ha insistido a lo largo del documento, en este trabajo se pretende analizar tanto la estructura cardiaca (segmentación ventricular), la información de movimiento (flujo óptico denso), así como también una versión integrada de las dos fuentes de información. Para la validación cuantitativa de los vectores embebidos se desarrollaron tres tipos de análisis: 1) cuantificación del espacio embebido con UMAP, 2) validación por clase utilizando las diferentes estrategias de clasificación y 3) valoración con diferente métricas para

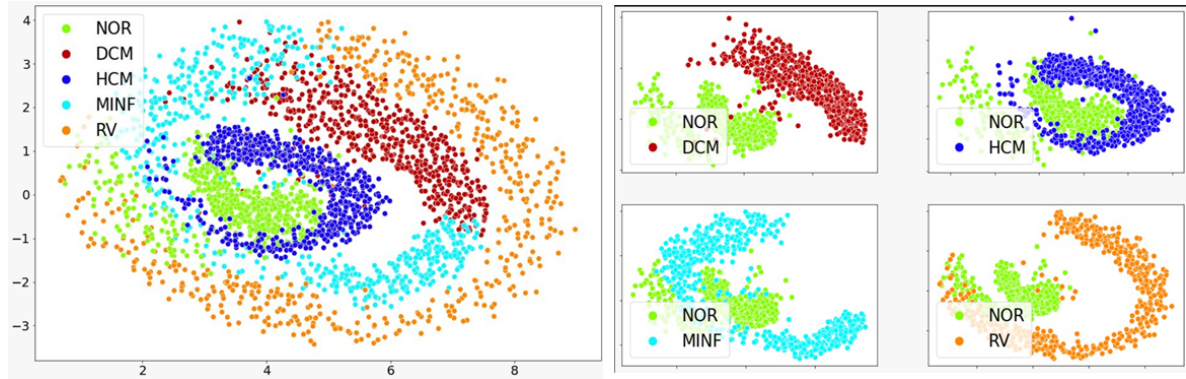
evaluar su capacidad discriminatoria entre patologías.

5.1. REPRESENTACIÓN PROFUNDA CONDICIONADA CON INFORMACIÓN MORFOLÓGICA: SEGMENTACIÓN

En un primer experimento se evaluó la capacidad de la arquitectura para representar condiciones cardiacas, usando vectores embebidos y utilizando como condición las segmentaciones ventriculares. Una vez entrenada la arquitectura, se generaron aproximadamente 563 muestras por cada condición. Con las muestras resultantes se retro-propagó sobre la arquitectura para obtener un vector embebido por cada imagen generada. Con los vectores embebidos resultantes se procedió a hacer el análisis cuantitativo, iniciando por el agrupamiento de estos vectores usando la estrategia de UMAP. En la figura 18 se puede observar un agrupamiento resultante (entre las patologías cardiacas) de los embebidos sacados de la GAN entrenada con el *priori* de la segmentación más la patología asociada, aunque también se evidencia una región donde se cruzan las patologías con condición Normal (NOR), Infarto en el miocardio previo (MINF) y Ventrículo anormal derecho (RV). Como se puede observar, la representación topológica obtenida con UMAP resulta sobresaliente, logrando separar las diferentes condiciones cardiacas, lo que indica una apropiada proyección de la representación de las muestras obtenidas. Los experimentos con UMAP se realizaron tanto de forma multi-clase (con todas las condiciones), como en términos binarios, siendo ambas satisfactorias para la tarea propuesta.

En una segunda validación se procedió a tomar los vectores de baja dimensionalidad y realizar un proceso de clasificación supervisada. Para ello se utilizaron un total de tres estrategias de clasificación y se obtuvieron las matrices de confusión correspondientes, las cuales dan información sobre la distribución de las predicciones obtenidas. En la figura 19 se pueden observar los resultados obtenidos, siendo sobresaliente la estrategia KNN reportando una sensibilidad promedio del 89.37 %.

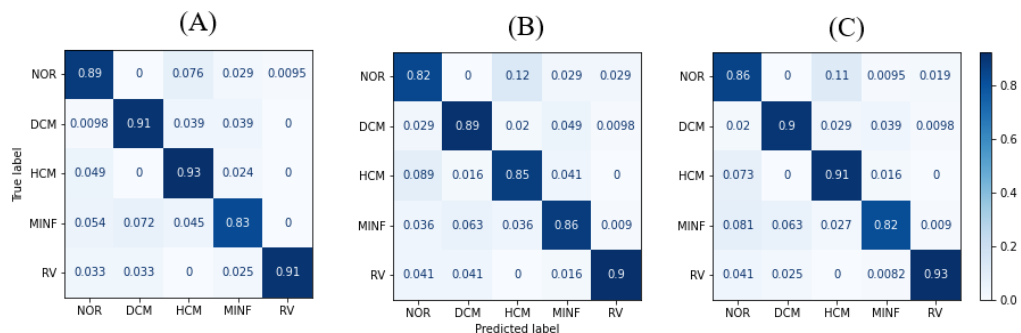
Figura 18. Agrupamiento no supervisado UMAP de los vectores embebidos condicionados. Se utilizó el *priori* de segmentación en este experimento.



Fuente: elaboración propia.

Estos resultados resultan fundamentales para mostrar la capacidad de los vectores embebidos en cuanto a la capacidad de recuperar medidas morfológicas discriminantes entre diferentes patologías, teniendo además un apropiado comportamiento con los diferentes clasificadores. Una de las principales limitaciones en la representación está relacionada con la condición RV (ventrículo derecho anormal), lo cual es consistente, teniendo en cuenta que la condición en la arquitectura se enfoca principalmente en el ventrículo izquierdo.

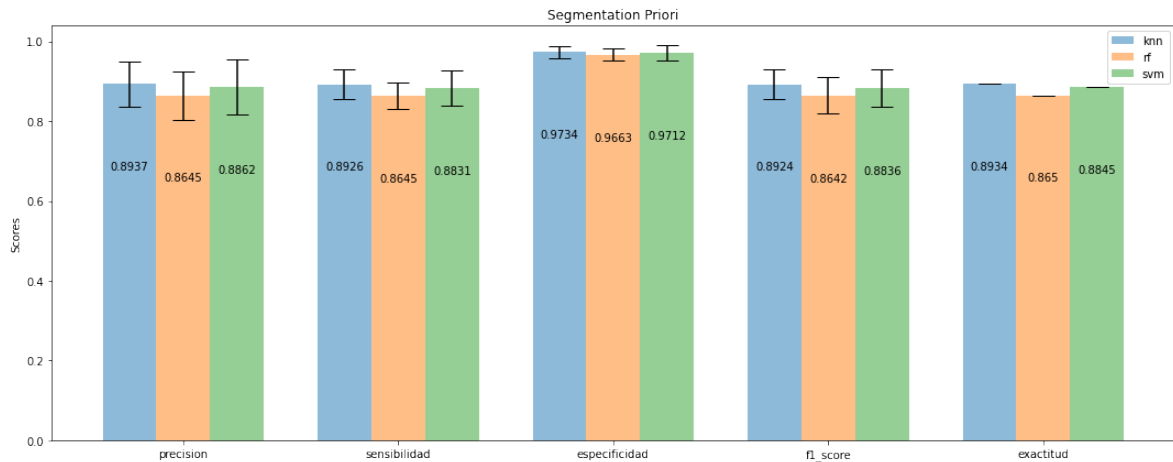
Figura 19. (A) matriz de confusión con el modelo KNN, (B) matriz de confusión con el modelo Random Forest, (C) matriz de confusión con el modelo support vector machine.



Fuente: elaboración propia.

En la figura 20 se presenta un diagrama de barras con diferentes métricas resultantes que complementan el análisis realizado sobre la arquitectura configurada con la condición de segmentación. Como es de esperarse, para un dataset relativamente balanceado, los tres clasificadores presentan puntajes relevantes relacionados con la precisión, sensibilidad, especificidad e incluso con respecto al *F1-score*. El modelo SVC, tiene una mayor desviación estándar (mostrada como una línea vertical en el borde superior de cada barra) comparado con los otros modelos. También que su exactitud junto con el modelo KNN es mayor a 88%. Cabe también resaltar la capacidad de especificidad del método, teniendo los tres métodos puntajes mayores al 96%, lo cual resulta primordial en las aplicaciones clínicas para el soporte de los análisis clínicos.

Figura 20. Diagrama de barras para el *priori* de segmentación

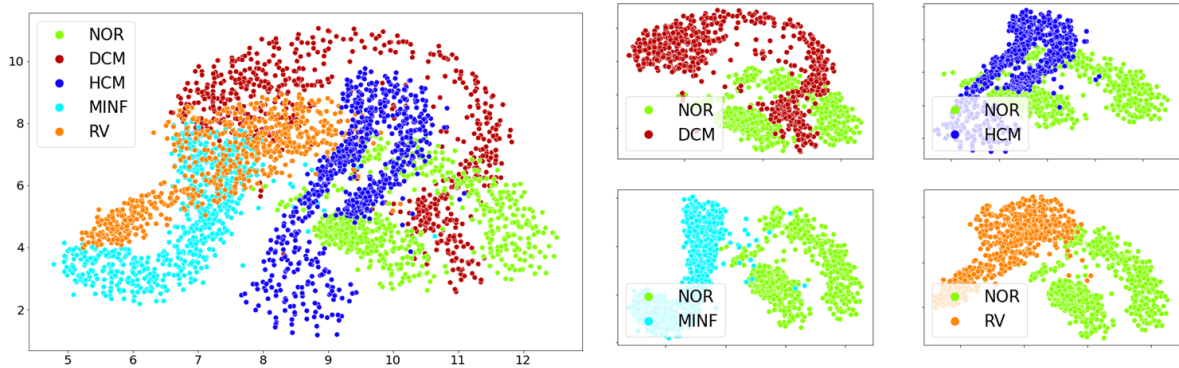


Fuente: elaboración propia.

5.2. REPRESENTACIÓN PROFUNDA CONDICIONADA CON INFORMACIÓN DE MOVIMIENTO: FLUJO ÓPTICO

En una segunda validación experimental se procedió a analizar la información dinámica cardíaca como información contrastiva para fijar una arquitectura de representación cardíaca. Se procedió entonces a evaluar la capacidad de la arquitectura para representar patologías, usando los vectores embebidos resultantes del entrenamiento con el flujo óptico denso. En este caso, una vez se entrenó la arquitectura, se generaron aproximadamente 672 muestras de cada clase, con el fin de obtener los vectores embebidos que fueron analizados por medio de un *clustering* a partir de una estrategia de UMAP. Los resultados obtenidos, utilizando la proyección a una dimensionalidad baja se pueden observar en la Figura 21. Se puede evidenciar que el agrupamiento resulta un poco más desafiante por la intersección y solapamiento entre algunas muestras de patologías diferentes. Esto comparado con el claro agrupamiento mostrado durante la segmentación. Sin embargo, la mayoría de los puntos son consistentes con un grupo característico, asociado a cada una de las patologías evaluadas. Del mismo modo, cuando se compara de manera binaria, se evidencia una mejor clasificación de clases, aunque este criterio no es suficiente para comprobar la representación de los embebidos con su respectiva patología.

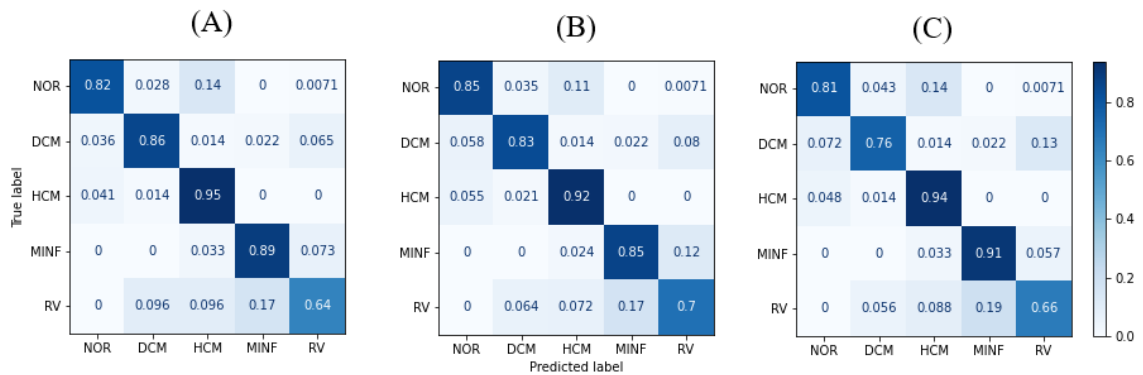
Figura 21. Agrupamiento no supervisado UMAP de los vectores embebidos condicionados. Se utilizó el *priori* del flujo óptico en este experimento.



Fuente: elaboración propia.

Siguiendo el mismo esquema metodológico de evaluación, se procedió entonces a hacer una segunda validación, por medio de los vectores embebidos de baja dimensionalidad resultantes de la estrategia de UMAP. Se usaron las mismas tres estrategias de clasificación supervisada, y el resultado de estas se pudo evidenciar en la figura 22, donde se obtuvo una sensibilidad promedio mayor a un 80 % para los tres modelos de clasificación, siendo el modelo KNN el mejor de nuevo, con una sensibilidad promedio de 83,29%. También se logra evidenciar que la sensibilidad de la patología de ventrículo derecho anormal vuelve a ser la más deficiente en los tres modelos, esta patología es confundida por los modelos de clasificación, con la patología de infarto en el miocardio previo (MINF). Este hecho puede estar asociado debido a que la fracción de eyección en estas patologías es similar e inferior al rango normal.

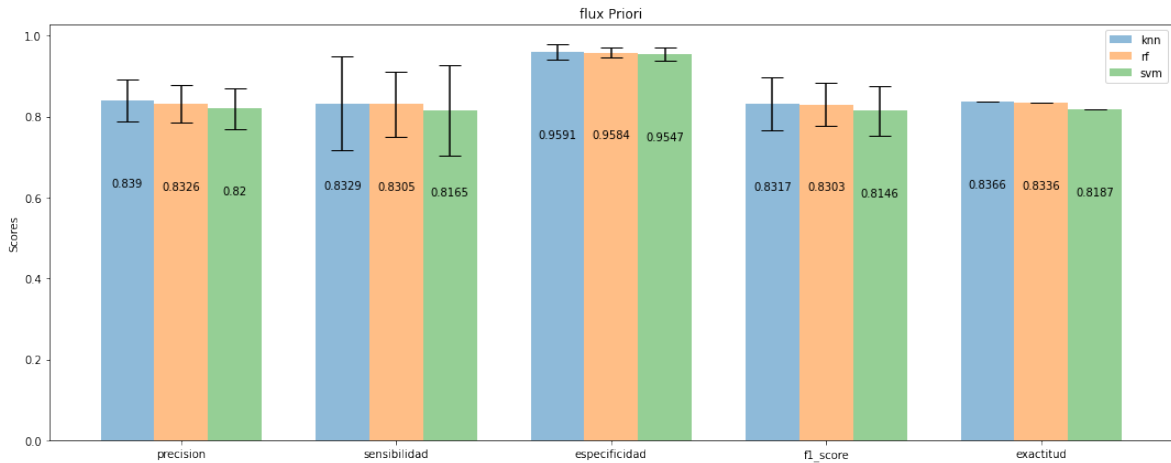
Figura 22. (A) matriz de confusión con el modelo KNN, (B) matriz de confusión con el modelo Random Forest, (C) matriz de confusión con el modelo support vector machine.



Fuente: elaboración propia.

De la figura 23, la cual representa un diagrama de barras con las diferentes métricas resultantes de la clasificación con los tres modelos entrenados anteriormente, se puede observar que la sensibilidad presenta un rango de desviación (también representado como una línea vertical en el borde superior de cada barra) considerablemente alto para los tres casos, esto debido al problema mencionado anteriormente con las patologías MINF y RV debido a su similar capacidad para bombear sangre. Aún así, la capacidad de especificidad sigue siendo sobresaliente con valores cercanos a un 95 %, lo cual resalta confiabilidad a la hora de ser usados como soporte en los análisis clínico.

Figura 23. Diagrama de barras para el *priori* de flujo óptico

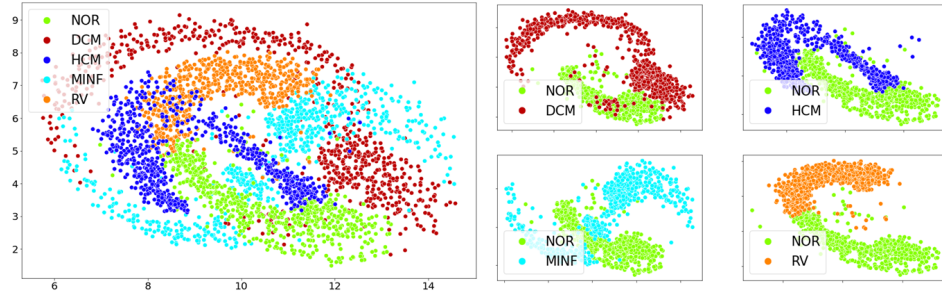


Fuente: elaboración propia.

5.3. REPRESENTACIÓN PROFUNDA CONDICIONADA CON INFORMACIÓN MORFOFISIOLÓGICA INTEGRADA

Teniendo en cuenta los resultados anteriores, y para evaluar una representación que integre tanto información estructural como información de movimiento, en un tercer experimento se utilizó el *priori* del flujo óptico únicamente para el área cardiaca junto con su patología asociada, obteniendo como resultado vectores embebidos que marcan buenos patrones de agrupamiento con respecto a las diferentes patologías como se puede evidenciar en la figura 24. Como se muestra para este experimento, los vectores referentes a la patología de infarto en el miocardio previo (MINF) se cruzan a lo largo de las otras cuatro patologías, aún así, si comparamos de manera binaria las otras con respecto a la normal, se puede observar una óptima separación de clases, indicando una mayor eficiencia en la representación de patologías cardiacas.

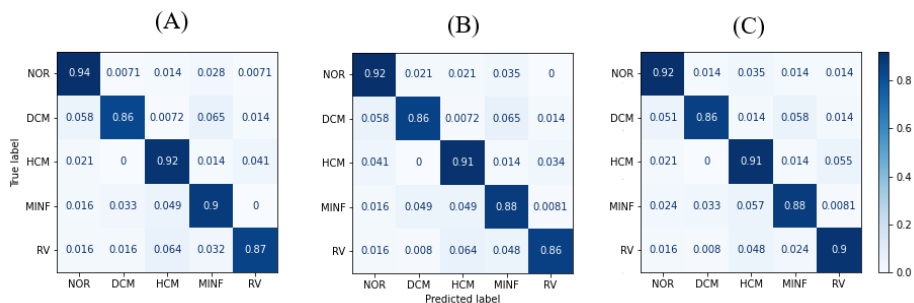
Figura 24. Agrupamiento no supervisado UMAP de los vectores embebidos condicionados. Se utilizaron los aprioris de segmentación y flujo óptico en este experimento.



Fuente: elaboración propia.

Respecto a la clasificación de embebidos resultantes de la reducción de dimensionalidad UMAP con la combinación de los *prioris* anteriores, se utilizaron los mismos tres modelos para comparar la eficiencia de combinar estos dos *prioris*. En la figura 25, en general para los tres modelos, se pudo evidenciar una mejoría para las patologías con condición Normal (NOR) e infarto en el miocardio previo (MINF). Además, el usar este *priori* integrado remueve la limitante de la patología de ventrículo derecho anormal (RV), ya que también mejoró con respecto al flujo óptico visible en la figura 22. En general, la sensibilidad promedio supera el 88 % lo cual denota resultados satisfactorios para la tarea de clasificación.

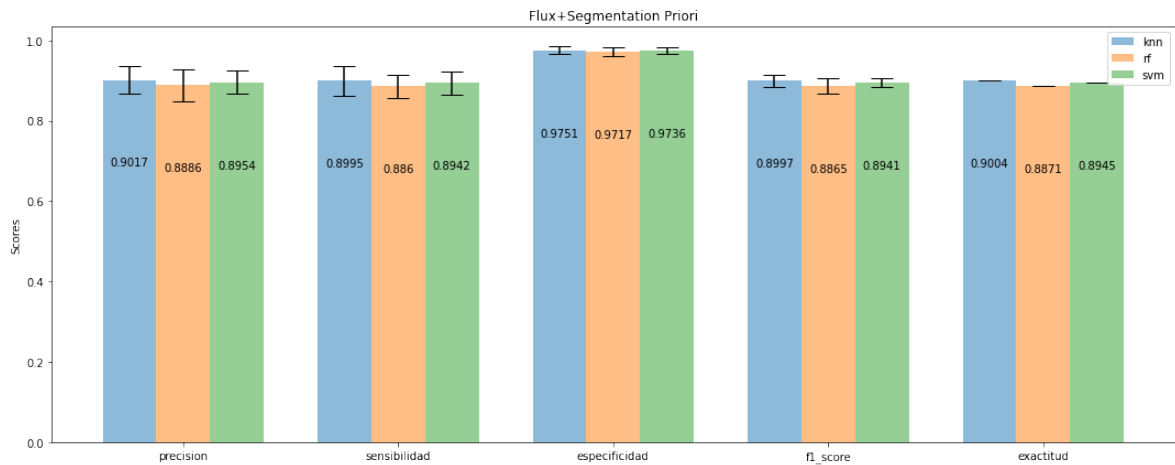
Figura 25. (A) matriz de confusión con el modelo KNN, (B) matriz de confusión con el modelo Random Forest, (C) matriz de confusión con el modelo support vector machine.



Fuente: elaboración propia.

En el diagrama de barras de la figura 26 se pudo comprobar que las métricas son mejores respecto a las de únicamente el flujo, como se puede ver en la figura 23, también, que el rango de su desviación se ve disminuido, lo cual refleja una mayor precisión en la tarea de clasificación. Con lo anterior, se puede concluir que estos embebidos son los que mejor se desempeñan en tareas de clasificación supervisada, con una exactitud del 90,04 % para el modelo de KNN, junto con una especificidad del 97,51 %, siendo este modelo, el que mejor resultado ha generado del conjunto de experimentos en general, modelo el cual sirve como soporte para afirmar que el uso de estos embebidos serviría de ayuda diagnóstica para descartar patologías cardíacas, y en general, para ayudar en tareas de decisión cuando se refiere al diagnóstico de patologías del corazón (sólo si consideramos únicamente las patologías con las que fue entrenada la arquitectura).

Figura 26. Diagrama de barras para el *priori* del flujo óptico segmentado en el corazón más la patología



Fuente: elaboración propia.

6. CONCLUSIONES Y PERSPECTIVAS

En este trabajo se implementó una representación generativa convolucional de imágenes cine-MRI para evaluar la capacidad de codificar patrones cardíacos con carácter discriminatorio entre patologías. Durante el desarrollo del trabajo se implementó una arquitectura generativa múltiple condicional soportada por *prioris* cardíacos con el fin de aprender mejores características relativas a múltiples patologías asociadas, y así poder obtener nuevos biomarcadores cardíacos. La arquitectura desarrollada tiene la capacidad de incluir información relacionada con diagnósticos del volumen cardíaco, información espacial relacionada con la segmentación del ventrículo izquierdo, así como también información relativa al movimiento cardíaco, implementando un flujo óptico denso de largos desplazamientos. Durante el entrenamiento y evaluación se consideró un conjunto de volúmenes cine-MRI, donde se extrajeron imágenes principalmente basales, las cuales permitieron entrenar la red, así como también validar su capacidad de correspondencia con diferentes patologías. Para la validación de un descriptor cardíaco, se calcularon vectores embebidos a partir de las imágenes generadas, utilizando un esquema de retro-propagación que afinaba el vector de entrada en el segmento generativo de la red. Los vectores embebidos resultantes fueron proyectados a estrategias de clasificación para evidenciar y validar la capacidad discriminatoria entre diferentes patologías de entrenamiento.

Se realizaron diferentes experimentos que permitieran evaluar la capacidad e importancia de restricciones tanto estructurales (segmentación) como morfológicas (flujo óptico). En primera instancia, las imágenes generadas por la segmentación tienen una mayor calidad observacional, se ajustan en mayor medida a las estructuras cardíacas típicas, pero su generación es transversal entre clases. En este sentido,

en la evaluación de clasificación estas medidas estructurales reportan una mayor limitación. En cuanto a las imágenes generadas, utilizando como restricción los patrones de movimiento, codificados como flujo óptico, se puede ver una desmejora en la representación visual. Esto debido a que los patrones de flujo no tienen una representación espacial precisa y hacen referencia a la dinámica entre varias imágenes secuenciales. Sin embargo, los embebidos resultantes de esta representación tienen un mayor carácter discriminatorio y las métricas reportadas muestran un potencial uso en la rutina clínica como soporte para la caracterización de las patologías. Se evidenció, después del proceso de validación de estos espacios embebidos, que las condiciones del flujo óptico segmentado más su patología asociada produce embebidos que resultan más sencillos de separar para clasificadores de inteligencia artificial, como se puede observar en la figura 25. En efecto la integración tanto de información estructural como de movimiento puede aportar en mayor medida a la caracterización de patologías, así como también a la búsqueda de nuevas relaciones que permitan una mayor diferenciación entre condiciones cardiacas con descripciones diagnósticas similares.

Las múltiples condiciones en las arquitecturas generativas pueden ser clave para su uso en la rutina clínica, siendo una fuente de información primordial para la caracterización y generación de volúmenes cardiacos. También para el uso explícitamente generativo, observacional y de aumento de datos, se recomienda usar condiciones espaciales, con referencias morfológicas. En cuanto a la cuantificación de nuevos patrones, las arquitecturas que admiten múltiples condiciones pueden ser claves para ahondar en nuevas relaciones y patrones ocultos que puedan apoyar tareas de diagnóstico. A pesar de los resultados significativos y la capacidad de modelar patrones cardiacos complejos, estas arquitecturas siguen siendo dependientes del conjunto de datos y por supuesto de la capacidad computacional para operar sobre

ellos. Es por ello, que en futuros trabajos se recomiendan hacer adaptaciones de la arquitectura y a los esquemas de entrenamiento, por ejemplo utilizando secuencias progresivas que permitan mitigar el tiempo de cálculo para obtener una apropiada representación.

BIBLIOGRAFÍA

ABRAHAM, Gilu K.; JAYANTHI, V.S. y BHASKARAN, Preethi. “10 - Convolutional neural network for biomedical applications”. En: *Computational Intelligence and Its Applications in Healthcare*. Ed. por Jitendra Kumar VERMA; Sudip PAUL y Prashant JOHRI. Academic Press, 2020, págs. 145 -156. DOI: <https://doi.org/10.1016/B978-0-12-820604-1.00010-8> (vid. pág. 17).

ALFAKIH, Khaled, *et al.* “Assessment of ventricular function and mass by cardiac magnetic resonance imaging”. En: *European radiology* 14.10 (2004), págs. 1813-1822 (vid. pág. 37).

ALTINTAS, Zeynep; FAKANYA, Wellington M. y TOTHILL, Ibtisam E. “Cardiovascular disease detection using bio-sensing techniques”. En: *Talanta* 128 (2014), págs. 177 -186. DOI: <https://doi.org/10.1016/j.talanta.2014.04.060> (vid. pág. 23).

ARJOVSKY, Martin; CHINTALA, Soumith y BOTTOU, Léon. “Wasserstein gan”. En: *arXiv preprint arXiv:1701.07875* (2017) (vid. pág. 20).

AVENDI, M.R.; KHERADVAR, Arash y JAFARKHANI, Hamid. “A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI”. En: *Medical Image Analysis* 30 (2016), págs. 108 -119. DOI: <https://doi.org/10.1016/j.media.2016.01.005> (vid. págs. 26, 28).

BALAMURALI MURUGESAN Vijaya Raghavan S, Kaushik Sarveswaran Keerthi Ram Mohanasankar Sivaprakasam. “A Global-Local context based Generative Adversarial Network for MRI Reconstruction”. En: *arXiv preprint arXiv:1908.09262* (2019) (vid. pág. 19).

BERMUDEZ, Camilo, *et al.* “Learning implicit brain MRI manifolds with deep learning”. En: *Medical Imaging 2018: Image Processing*. Vol. 10574. International Society for Optics y Photonics. 2018, pág. 105741L (vid. pág. 20).

BIFFI, C., *et al.* “3D High-Resolution Cardiac Segmentation Reconstruction From 2D Views Using Conditional Variational Autoencoders”. En: *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. 2019, págs. 1643-1646. DOI: 10.1109/ISBI.2019.8759328 (vid. pág. 37).

BONOW RO Fonarow GC, O’Gara PT Yancy CW. *Association of Coronavirus Disease 2019 (COVID-19) With Myocardial Injury and Mortality*. 2020 (vid. pág. 11).

BORLAND, Lisa y HEEL, Marin van. “Classification of image data in conjugate representation spaces”. En: *J. Opt. Soc. Am. A* 7.4 (1990), págs. 601-610. DOI: 10.1364/JOSAA.7.000601 (vid. pág. 25).

BROX, T.; BREGLER, C. y MALIK, J. “Large displacement optical flow”. En: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 2009, págs. 41-48. DOI: 10.1109/CVPR.2009.5206697 (vid. pág. 40).

CALIMERI, Francesco, *et al.* “Biomedical data augmentation using generative adversarial neural networks”. En: *International conference on artificial neural networks*. Springer. 2017, págs. 626-634 (vid. pág. 19).

CHAPARRO-MUÑOZ, Marinela y PRASAD, Sanjay K. “Valor diagnóstico de la resonancia magnética cardiaca en la insuficiencia cardiaca: estado actual”. En: *Cardio-core* 46.1 (2011), págs. 4-12 (vid. pág. 29).

CHARTSIAS, Agisilaos, *et al.* “Adversarial image synthesis for unpaired multi-modal cardiac data”. En: *International Workshop on Simulation and Synthesis in Medical Imaging*. Springer. 2017, págs. 3-13 (vid. pág. 22).

EMAMI, Hajar, *et al.* “Generating synthetic CTs from magnetic resonance images using generative adversarial networks”. En: *Medical physics* 45.8 (2018), págs. 3627-3636 (vid. pág. 21).

GIHYUN KWON Chihye Han, Dae-shik Kim. “Generation of 3D Brain MRI Using Auto-Encoding Generative Adversarial Networks”. En: *arXiv preprint arXiv:1908.02498* (2019) (vid. pág. 20).

GOKSEL, D.; OZKAN, M. y OZTURK, C. “Cardiac motion analysis in MRI for classification”. En: *Proceedings IEEE International Symposium on Biomedical Imaging*. 2002, págs. 935-938. DOI: 10.1109/ISBI.2002.1029415 (vid. pág. 40).

GOODFELLOW, Ian. “NIPS 2016 tutorial: Generative adversarial networks”. En: *arXiv preprint arXiv:1701.00160* (2016) (vid. pág. 18).

GOODFELLOW, Ian, *et al.* “Generative adversarial nets”. En: *Advances in neural information processing systems*. 2014, págs. 2672-2680 (vid. pág. 17).

HAN, Changhee, *et al.* “GAN-based synthetic brain MR image generation”. En: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE. 2018, págs. 734-738 (vid. pág. 19).

ISIN, Ali y OZDALILI, Selen. “Cardiac arrhythmia detection using deep learning”. En: *Procedia Computer Science* 120 (2017). 9th International Conference on Theory and Application of Soft Computing, Computing with Words and Perception, ICSCCW

2017, 22-23 August 2017, Budapest, Hungary, págs. 268 -275. DOI: <https://doi.org/10.1016/j.procs.2017.11.238> (vid. pág. 27).

JIN, Cheng-Bin, *et al.* “Deep CT to MR synthesis using paired and unpaired data”. En: *Sensors* 19.10 (2019), pág. 2361 (vid. págs. 21, 22).

KANG, Dongwoo, *et al.* “Heart chambers and whole heart segmentation techniques: review”. En: *Journal of Electronic Imaging* 21.1 (2012), págs. 1 -17. DOI: 10.1117/1.JEI.21.1.010901 (vid. pág. 37).

KARRAS, Tero, *et al.* “Progressive growing of gans for improved quality, stability, and variation”. En: *arXiv preprint arXiv:1710.10196* (2017) (vid. pág. 53).

KINGMA, Diederik P. y WELLING, Max. “An Introduction to Variational Autoencoders”. En: *Foundations and Trends® in Machine Learning* 12.4 (2019), 307–392. DOI: 10.1561/22000000056 (vid. pág. 24).

KONONENKO, Igor. “Machine learning for medical diagnosis: history, state of the art and perspective”. En: *Artificial Intelligence in medicine* 23.1 (2001), págs. 89-109 (vid. pág. 12).

LI, Ziqiang, *et al.* *Interpreting the Latent Space of GANs via Correlation Analysis for Controllable Concept Manipulation*. 2020. arXiv: 2006.10132 [cs.CV] (vid. pág. 13).

MAKHZANI, Alireza, *et al.* *Adversarial Autoencoders*. 2016. arXiv: 1511.05644 [cs.LG] (vid. pág. 24).

MAN, R.; YANG, P. y XU, B. “Classification of Breast Cancer Histopathological Images Using Discriminative Patches Screened by Generative Adversarial Networks”.

En: *IEEE Access* 8 (2020), págs. 155362-155377. DOI: 10.1109/ACCESS.2020.3019327 (vid. pág. 14).

MCINNES, Leland; HEALY, John y MELVILLE, James. *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction*. 2020. arXiv: 1802.03426 [stat.ML] (vid. pág. 46).

MEDINA, Ruben, *et al.* "Cardiac Motion Estimation in Magnetic Resonance Images Using Optical Flow". En: *IEEE Latin America Transactions* 14 (jun. de 2016), págs. 2807-2816. DOI: 10.1109/TLA.2016.7555259 (vid. pág. 40).

MENDIS, Shanthi, *et al.* *Global atlas on cardiovascular disease prevention and control*. World Health Organization, 2011 (vid. pág. 11).

MOK, Tony CW y CHUNG, Albert CS. "Learning data augmentation for brain tumor segmentation with coarse-to-fine generative adversarial networks". En: *International MICCAI Brainlesion Workshop*. Springer. 2018, págs. 70-80 (vid. pág. 20).

MUKHERJEE, Sudipto, *et al.* "ClusterGAN: Latent Space Clustering in Generative Adversarial Networks". En: *Proceedings of the AAAI Conference on Artificial Intelligence* 33.01 (2019), págs. 4610-4617. DOI: 10.1609/aaai.v33i01.33014610 (vid. pág. 26).

MURUGESAN, Balamurali, *et al.* "Recon-GLGAN: A Global-Local context based Generative Adversarial Network for MRI Reconstruction". En: (ago. de 2019) (vid. pág. 28).

NIE, Dong, *et al.* "Medical image synthesis with context-aware generative adversarial networks". En: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2017, págs. 417-425 (vid. pág. 21).

PEREZ, Luis y WANG, Jason. “The effectiveness of data augmentation in image classification using deep learning”. En: *arXiv preprint arXiv:1712.04621* (2017) (vid. pág. 13).

PUTIN, Evgeny, *et al.* “Deep biomarkers of human aging: application of deep neural networks to biomarker development”. En: *Aging (Albany NY)* 8.5 (2016), pág. 1021 (vid. págs. 24, 25).

RADFORD, Alec; METZ, Luke y CHINTALA, Soumith. “Unsupervised representation learning with deep convolutional generative adversarial networks”. En: *arXiv preprint arXiv:1511.06434* (2015) (vid. pág. 19).

RONNEBERGER, Olaf; FISCHER, Philipp y BROX, Thomas. “U-Net: Convolutional Networks for Biomedical Image Segmentation”. En: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. por Nassir NAVAB, *et al.* Cham: Springer International Publishing, 2015, págs. 234-241 (vid. pág. 37).

SANTOS TANAKA, Fabio Henrique Kiyoiiti dos y ARANHA, Claus. *Data Augmentation Using GANs*. 2019. arXiv: 1904.09135 [cs.LG] (vid. pág. 32).

SAVADJIEV, Peter, *et al.* “Demystification of AI-driven medical image interpretation: past, present and future”. En: *European Radiology* 29 (3 mar. de 2019), págs. 1616-1624. DOI: 10.1007/s00330-018-5674-x (vid. pág. 23).

SHARMA, Palash. *Autoencoders in Keras – Introduction to Beginners with Example*. URL: <https://machinelearningknowledge.ai/autoencoders-in-keras-introduction-to-beginners-with-example/> (vid. pág. 18).

VELÁZQUEZ, Edwin; BRAVO, Antonio y VERA, Miguel. "Segmentación del ventrículo izquierdo en imágenes cardíacas usando técnicas no paramétricas". En: mayo de 2015. DOI: 10.13140/RG.2.1.1910.2888 (vid. pág. 37).

VOYNOV, Andrey y BABENKO, Artem. *Unsupervised Discovery of Interpretable Directions in the GAN Latent Space*. 2020. arXiv: 2002.03754 [cs.LG] (vid. pág. 26).

WATANABE, H., *et al.* "Bone Metastatic Tumor Detection based on AnoGAN Using CT Images". En: *2019 IEEE 1st Global Conference on Life Sciences and Technologies (LifeTech)*. 2019, págs. 235-236. DOI: 10.1109/LifeTech.2019.8883999 (vid. pág. 14).

WONG, Sebastien C, *et al.* "Understanding data augmentation for classification: when to warp?" En: *2016 international conference on digital image computing: techniques and applications (DICTA)*. IEEE. 2016, págs. 1-6 (vid. pág. 13).

WU, Jiajun, *et al.* "Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling". En: *Advances in Neural Information Processing Systems*. Ed. por D. LEE, *et al.* Vol. 29. Curran Associates, Inc., 2016, págs. 82-90 (vid. pág. 26).

YI, Xin; WALIA, Ekta y BABYN, Paul. "Generative adversarial network in medical imaging: a review". En: *arXiv preprint arXiv:1809.07294* (2018) (vid. págs. 16, 19).

YILDIRIM, Gökhan; SEWARD, Calvin y BERGMANN, Urs. *Disentangling Multiple Conditional Inputs in GANs*. 2018. arXiv: 1806.07819 [cs.CV] (vid. pág. 33).

ZABALZA, Jaime, *et al.* "Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging". En: *Neurocomputing* 185 (2016), págs. 1-10 (vid. pág. 24).

ZHANG, Le; GOOYA, Ali y FRANGI, Alejandro F. "Semi-supervised assessment of incomplete LV coverage in cardiac MRI using generative adversarial nets". En: *International Workshop on Simulation and Synthesis in Medical Imaging*. Springer. 2017, págs. 61-68 (vid. pág. 21).

ZHANG, Zizhao; YANG, Lin y ZHENG, Yefeng. "Translating and segmenting multimodal medical volumes with cycle-and shape-consistency generative adversarial network". En: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018, págs. 9242-9251 (vid. pág. 22).

ZHOU, S Kevin; GREENSPAN, Hayit y SHEN, Dinggang. *Deep learning for medical image analysis*. Academic Press, 2017 (vid. pág. 16).