

Identificación de ventrículo izquierdo del corazón mediante máscaras y landmarks utilizando modelos Deep Learning

Victor Hugo Arreola Elenes^[A01635682], Fernando Ojeda Marín^[A01639252], Mario Alberto Castañeda Martínez^[A01640152], and Luis Manuel Orozco Yáñez^[A01707822]

Instituto Tecnológico y de Estudios Superiores de Monterrey Campus Guadalajara

Resumen The Advanced AI Concentration for Data Science II project aims to implement and compare two segmentation approaches for the left ventricle of the heart: one based on binary masks and the other utilizing landmarks, using the EchoNet-Dynamic dataset. With 10,030 echocardiogram videos, the dataset is divided into training, validation, and test sets, each with frames labeled for left ventricle coordinates.

The more efficient method, identified through comparison, is integrated into a medical application system, employing Deep Learning algorithms and a U-Net architecture for image segmentation. The project, beyond implementation, seeks to select the most efficient method based on accuracy and processing speed, enhancing cardiac anomaly detection through left ventricle analysis.

This research improves efficiency in cardiac diagnosis, enabling informed decisions and personalized treatments. Advanced AI in medicine emphasizes innovation and continuous improvement. Echocardiograms (ECG) play a crucial role in evaluating heart activity, particularly in emergencies.

Understanding the left ventricle's state assists in distinguishing cardiac etiologies and prevalent heart failure cases in emergency departments. Bedside echocardiography offers a rapid, non-invasive method to assess systolic cardiac function.

Results show promising outcomes, especially in the mask model, with a Dice Score of approximately 0.91 on the test set. Consistency across metrics suggests model robustness. The landmark model demonstrates satisfactory fit, with some limitations indicating areas for improvement. Quantitative evaluation supports qualitative observations, marking success in left ventricle segmentation with potential for future enhancements and clinical applications.

Keywords: Deep Learning · Left Ventricle · Mask · Landmarks · Ecocardiogram

1. Introducción

El proyecto de la concentración en Inteligencia Artificial Avanzada para la Ciencia de Datos II consiste en implementar y comparar el rendimiento de dos enfoques, uno basado en máscaras binarias y el otro en máscaras usando puntos de referencia (landmarks) para la segmentación automática del ventrículo izquierdo del corazón haciendo uso del dataset *EchoNet-Dinamic* de la Universidad de Stanford. Este dataset contiene 10,030 videos de ecocardiogramas divididos previamente en conjuntos de entrenamiento, validación y prueba.

Adicionalmente, cada uno de los videos cuenta con 2 frames etiquetados con las respectivas coordenadas que delimiten el área del ventrículo izquierdo. Es importante mencionar que se tuvo un socio formador (Exo Imaging) el cual fue clave en la realización y guía en el reto, además se mantuvo una comunicación asertiva (sesiones de preguntas y correo electrónico) y también se tuvo apoyo con la obtención de los datos (explicado más a detalle en la sección de aspectos éticos).

Hecha la comparativa entre ambos enfoques, se selecciona el más eficiente para ser implementado en un sistema de aplicación médica que agilice la detección y generación de diagnósticos a partir de las capacidades del ventrículo izquierdo según su tamaño estimado. Para esta tarea, se utilizaron algoritmos de Aprendizaje Profundo (Deep Learning) junto a una arquitectura U-Net para la segmentación de las imágenes obtenidas a partir de los videos.

Por ende el proyecto no solo se limita a la implementación y comparación de los dos enfoques propuestos, sino que también busca la selección del método más eficiente. El criterio de eficiencia se establece mediante la evaluación del rendimiento de las estrategias en términos de precisión y velocidad de procesamiento. La estrategia seleccionada será integrada en un sistema de aplicación médica que tiene como objetivo acelerar y mejorar la detección de anomalías cardíacas a través del análisis del ventrículo izquierdo.

El impacto potencial de esta investigación se proyecta hacia la mejora de la eficiencia y precisión en el diagnóstico cardíaco, permitiendo a los profesionales de la salud tomar decisiones informadas y brindar tratamientos más personalizados. La aplicación de técnicas avanzadas de Inteligencia Artificial en el ámbito médico demuestra el compromiso con la innovación y el mejoramiento continuo de las prácticas clínicas para el beneficio de la salud humana.

En este contexto, se destaca la relevancia de las herramientas médicas avanzadas, como lo son los ecocardiogramas (ECG), herramientas fundamentales en medicina para evaluar la actividad cardíaca. Estos registros gráficos proporcionan información crucial sobre la regularidad del ritmo cardíaco y la presencia de posibles anomalías cardíacas. En situaciones de emergencia, los ECG desempeñan un papel vital al permitir una rápida identificación de condiciones cardíacas agudas, como el infarto de miocardio, arritmias graves o bloqueos cardíacos.

Conocer el estado del ventrículo izquierdo del corazón a través de ecocardiogramas puede ayudar a distinguir entre etiologías cardíacas y otras causas de hipotensión no diferenciada. En la actualidad, la insuficiencia cardíaca congestiva es una afección prevalente en la población del departamento de emergencias, en donde más de la mitad de los pacientes con disfunción sistólica moderada a grave no han sido diagnosticados con insuficiencia cardíaca.

La ecocardiografía realizada en la cabecera por el médico de emergencias ofrece un método rápido, no invasivo y económico para determinar el papel que la función cardíaca sistólica del paciente puede desempeñar en el proceso de la enfermedad. La evaluación de la función del ventrículo izquierdo (VI) puede ayudar a diagnosticar la insuficiencia cardíaca, identificar derrames pericárdicos, diagnosticar la tensión del ventrículo derecho, prever la respuesta a los líquidos y prever resultados a corto plazo en caso de paro cardíaco. Los

médicos de emergencias con formación específica en ecocardiografía transtorácica pueden determinar con precisión la fracción de eyección del ventrículo izquierdo (FEVI) en pacientes críticamente enfermos. (Now, 2015)

2. Metodología

2.1. Exploración de las bases de datos

El dataset de *EchoNet-Dynamic* cuenta con tres diferentes bases de datos, siendo una de ellas una carpeta llamada "Videos" junto con dos archivos csv.

Dentro de la primer carpeta llamada "Videos" podemos encontrar 10,030 videos en formato .avi de ecocardiogramas con una duración de 3 segundos cada uno, los cuales muestran la actividad cardíaca desde una vista apical. Particularmente, se puede apreciar el momento de mayor y menor contracción en el movimiento del corazón.

Se tiene primero un archivo csv que, de entre diferentes columnas, principalmente muestra el nombre de cada uno de los videos junto con el conjunto al que pertenece; TRAIN, VAL o TEST (Entrenamiento, Validación o Prueba).

Complementariamente, se cuenta con otro archivo .csv que, muestra nuevamente el nombre de cada uno de los videos en la carpeta "Videos", junto con un conjunto de 20 coordenadas (X1, Y1, X2 y Y2) que juntas determinan la ubicación del ventrículo izquierdo del corazón. También se tienen 2 frames previamente etiquetados por video. Estos describen los 20 puntos de referencia en los puntos de mayor y menor contracción del corazón; los puntos donde más y menos sangre se bombea.

La manera en que se guardaron los datos, es que al tener acceso a un archivo Zip con las tres bases de datos. Se guardaron las bases de datos en Google Drive, con el objetivo de que todos los integrantes del equipo puedan acceder de forma sencilla a los datos y con esto, cada quien pueda trabajar o experimentar con los datos proveídos. Básicamente el uso de Drive fue crucial para el manejo de datos debido a que al ser un servicio en la nube, resultó más fácil el intercambio de información en el equipo. Es importante añadir que se realizó una generación de carpetas en Drive para el diferente manejo de datos o para guardar datos generados por el equipo (imágenes, máscaras, landmarks).

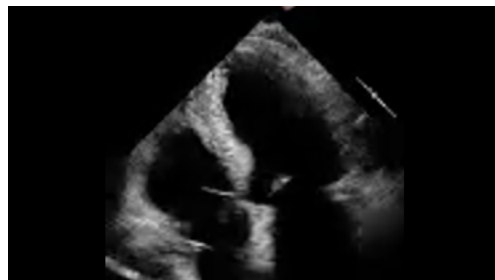


Figura 1: Muestra de apreciación de un Ecocardiograma en la base de datos.

La segunda base de datos después de la carpeta con los videos, consta de una descripción general de cada video, donde se incluyen variables como el nombre del video, el alto y ancho de sus frames y a qué conjunto pertenece el video; Entrenamiento, Validación o Prueba.

Finalmente, se cuenta con otra base de datos la cual contiene los dos frames específicos de cada video donde en un frame del video se aprecia el ventrículo en su punto máximo de dilatación, junto a otro frame con el ventrículo en su punto de máxima contracción. Para cada frame del video, se tienen 20 puntos o coordenadas (landmarks) que segmentan la ubicación del ventrículo izquierdo.

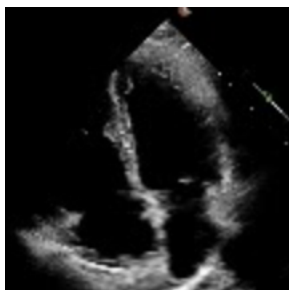


Figura 2: Frame del video con ventrículo izquierdo más dilatado

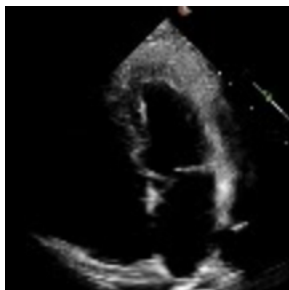


Figura 3: Frame del video con ventrículo izquierdo más contraído

En las Figuras 2 y 3 se muestran dos frames del mismo video. En la Figura 2, puede apreciarse la máxima amplitud y expansión del ventrículo durante una fase de relajación, facilitando así la entrada de sangre al ventrículo y su consiguiente llenado. Por otra parte, la Figura 3 exhibe una disminución en el tamaño del ventrículo durante una fase de contracción, momento en el cual el ventrículo impulsa la sangre oxigenada hacia la aorta.

En este momento, para cada video de la carpeta "Videos" se tienen identificados ambos frames anteriormente mencionados (máxima dilatación y contracción). Estos serán necesarios para generar nuevas imágenes a partir de las máscaras y coordenadas (landmarks) asociadas a cada uno de ellos.

Una máscara es un método utilizado para indicar qué elementos de una matriz deben ser utilizados y cuáles no deben serlo, esto significa que, en una matriz, una máscara binarizada vendría a tener 0's para indicar partes de la matriz que no se están utilizando para hacer una máscara y 1's para indicar las partes de la matriz donde se está identificando o aplicando una máscara. (Tutorialsonpoint, s.f.) (ML-science, s.f.) En este contexto, a cada frame en el que aparezca el ventrículo izquierdo, se le aplicará una máscara para definir una región de interés dentro del mismo, la cual será el ventrículo izquierdo y así, entrenar un modelo para identificar el ventrículo.



Figura 4: Máscara aplicada a la Figura 2, donde se enfoca solo el ventrículo izquierdo

Como se observa en la Figura 4, al aplicar una máscara al frame original, se obtiene una máscara binarizada donde la parte de interés es el propio ventrículo izquierdo.

Además de la utilización de máscaras, se implementará el uso de landmarks como técnica de identificación del ventrículo izquierdo. Un landmark viene a ser un punto de interés o de referencia dentro una imagen y sirve para identificar una parte específica dentro de una imagen. (Datarya, s.f.) En el caso del reto, la implementación de landmarks combinará ciertos aspectos y técnicas que se mencionarán más adelante, pero que pretenden que la imagen final para entrenar el modelo, tal como se muestra en la Figura 5.

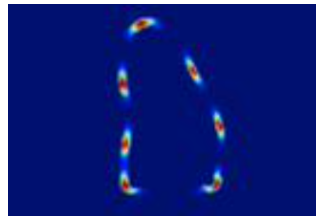


Figura 5: Aplicación de landmarks para la identificación del ventrículo izquierdo

Habiendo empezado con una exploración inicial de los datasets y una breve explicación sobre las técnicas a utilizar, se procederá a explicar los procesos realizados para la obtención de máscaras, landmarks y modelos de redes neuronales.

2.2. Obtención de las imágenes y máscaras

En este punto se cuenta con los datasets donde se muestran los dos frames específicos de cada video (mayor dilatación y contracción) junto con las coordenadas que segmentan el ventrículo izquierdo dentro de los mismos. A partir de aquí, se generarán nuevas imágenes por cada uno de los frames de los videos, para posteriormente trazar en ellos los puntos que generen las máscaras y los landmarks de cada uno.

	FileName	EF	ESV	EDV	FrameHeight	FrameWidth	FPS	NumberOfFrames	Split
0	0X100009310A3BD7FC	78.498406	14.881368	69.210534	112	112	50	174	VAL
1	0X1002E8FBACD08477	59.101988	40.383876	98.742884	112	112	50	215	TRAIN
2	0X1005D03EED19C65B	62.363798	14.267784	37.909734	112	112	50	104	TRAIN
3	0X10075961BC11C88E	54.545097	33.143084	72.914210	112	112	55	122	TRAIN
4	0X10094BA0A028EAC3	24.887742	127.581945	169.855024	112	112	52	207	VAL

Figura 6: Vistazo de base de datos donde se observa la columna "Split" que etiqueta los videos como train, val y test

Para generar la imágenes y máscaras, primero se generó un ambiente de Python a través de un archivo Jupyter Notebook, en donde se importaron bibliotecas útiles para el análisis de las bases de datos, como lo es la biblioteca Pandas. Para iniciar con la generación de las imágenes, se realizó una división en el dataset para tener 3 datasets distintos donde cada uno contenga los videos de entrenamiento, validación y prueba respectivamente.

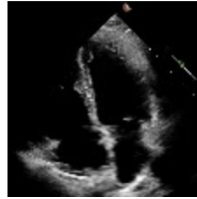
Posteriormente, se crea una conexión con la carpeta donde se encuentran los videos en formato .avi de los ecocardiogramas. En esta conexión se genera con la ruta de la carpeta y una lista de todos los nombres de los videos dentro de ella, esto con ayuda de la librería Os. Hecha esta conexión, es posible comenzar con la generación de una función para crear las imágenes de cada frame específico por cada video (2 frames), tal y como se muestra en la Figura 2 y la Figura 3.

Este proceso consiste en crear carpetas para almacenar imágenes y máscaras, seguido por la iteración a través de los videos dentro de una carpeta específica. Se verifican las coincidencias de nombres de videos entre la carpeta y un conjunto de datos para, posteriormente, extraer las coordenadas del ventrículo para cada combinación de video y frame. Se genera una máscara ordenando las coordenadas en sentido horario y se aplica sobre el frame, obteniendo como resultado una imagen donde resalta el ventrículo izquierdo.

Las máscaras y las imágenes originales se guardan en carpetas designadas con nombres identificadores. Este proceso estructurado facilita la obtención de información específica del ventrículo a partir de videos de ecocardiogramas.

La forma en que se muestran las imágenes junto con sus respectivas máscaras se muestra a continuación:

Frame 3 sin máscara:

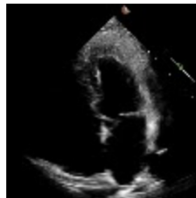


Frame 3 con máscara:



Figura 7: Imagen y máscara, en un frame específico con el ventrículo más dilatado.

Frame 18 sin máscara:



Frame 18 con máscara:



Figura 8: Imagen y máscara, en un frame específico con el ventrículo más contraído.

La Figura 7 y la Figura 8 muestran la generación de imágenes junto con sus respectivas máscaras para un mismo video. Cada uno de estos frames muestra al ventrículo izquierdo en su máximo punto de dilatación y contracción.

2.3. Generación de landmarks

Para generar los landmarks, primero se tienen que identificar los puntos que segmenten el ventrículo izquierdo dentro de la imagen. Este es un proceso automatizado para cada una de las imágenes dentro del conjunto de datos.

Ya con los landmarks identificados se procede a obtener el contorno de la máscara misma con una función de *find contours* de la biblioteca CV2, esto se hace a partir de ya tener las máscaras previamente binarizadas,

obteniendo el contorno de la misma en formato de arreglo *numpy*, esto es importante por que se quiere que la misma función de probabilidad de los landmarks, marque un contorno para tener landmarks más específicos, y lograr una función de probabilidad mejor adaptada ala forma del ventrículo.

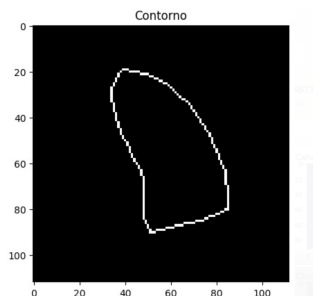


Figura 9: Contorno generado a partir de la máscara

Al momento de que se obtienen los *landmarks*, se le aplica una suavización mediante un kernel gaussiano para hacer una función de probabilidad con los mismos *landmarks*, con lo que se busca que es dar un cierto espacio donde se concentran los *landmarks*, siendo que justo el pixel donde se encuentra sea el valor más alto, que es de 1 y conforme se aleja de este punto gradualmente se va tomando valores menores, esto se hace con el fin de lograr que el modelo se adapte a una cierta zona donde los *landmarks* podrían estar, no castigando un valor de 0, por estar un pixel equivocado, se busca generar un modelo mejor adaptable y que conlleve una mejor generalización en otros ámbitos, en esta parte es de suma importancia barajar un valor de *sigma* adecuado, ya que un valor muy pequeño marcaría una zona muy específica, y uno muy alto haría que se pierda exactitud al marcar una región de probabilidad muy grande.

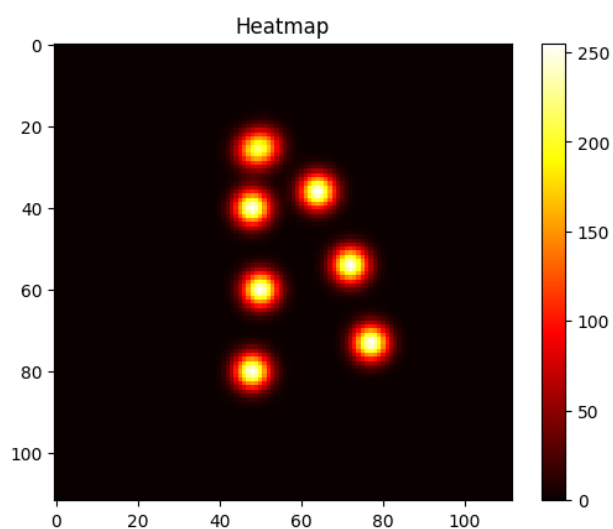


Figura 10: Landmarks Suavizados

Ya que se obtienen estos *landmarks* suavizados se pasa a multiplicar por el contorno de la máscara, para que la región de probabilidad sea aún más exacta, con esto se genera una mejor adaptación a lo que vendría siendo la segmentación del ventrículo izquierdo, así como de la misma manera para terminar con una forma en los datos de (112,112,7) teniendo 7 canales diferentes en el cual cada uno representa un *landmark* determinado, con lo que el modelo en cuestión pasa de ser una cuestión binaria y se transofrma en una cuestión multiclase, teniendo como cada clase un *landmark*.

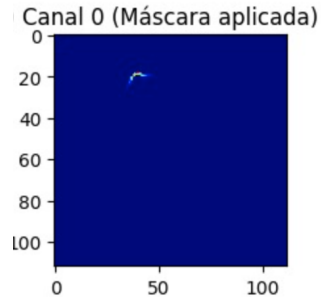


Figura 11: Canal 0 con máscara aplicada

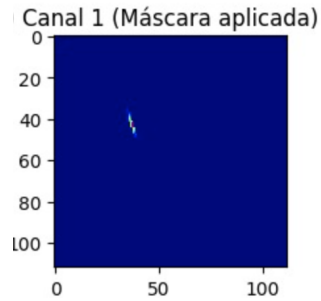


Figura 12: Canal 1 con máscara aplicada

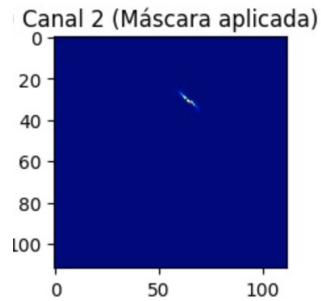


Figura 13: Canal 2 con máscara aplicada

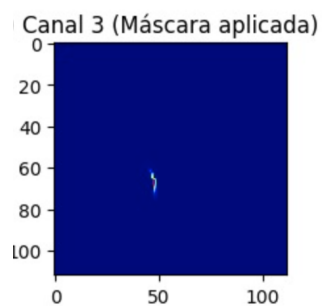


Figura 14: Canal 3 con máscara aplicada

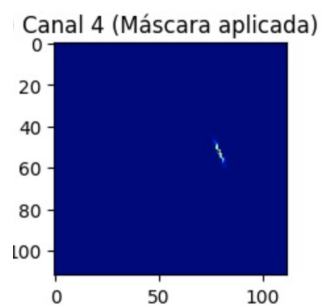


Figura 15: Canal 4 con máscara aplicada

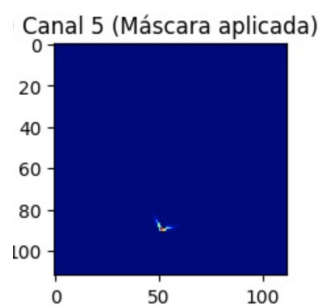


Figura 16: Canal 5 con máscara aplicada

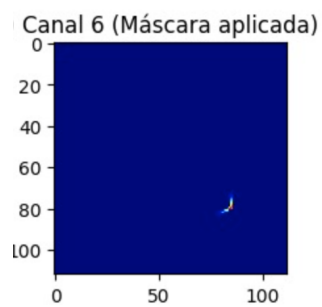


Figura 17: Canal 6 con máscara aplicada

Como se muestra en las imágenes anteriores, se obtienen siete canales para finalmente obtener una imagen general donde se identifica el ventrículo izquierdo del corazón, justamente a partir de una función de probabilidad, donde los valores más cercanos a 1 están justamente donde está el landmark y a partir de que se alejan mediante el contorno, dicha probabilidad va disminuyendo.

Aca se puede ver el mismo arreglo de landmarks, en escala de grises y alado de su imagen original respectiva, esto vendría siendo lo que se va a introducir al modelo en un principio.

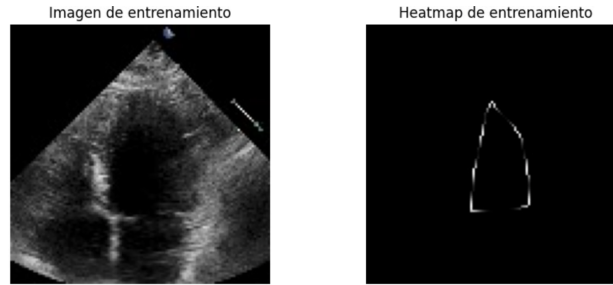


Figura 18: Ventrículo identificado usando los 7 canales

Así de la misma manera se automatizó a todo el modelo y se almacenaron los datos en formato de arreglos *numpy* para facilitar el proceso de generar y adaptar una red neuronal, adaptada a la problemática en cuestión.

2.4. Modelos utilizados

Para la realización del reto y la comparación de ambas técnicas para la identificación del ventrículo izquierdo en los ecocardiogramas, se realizó una aplicación de un modelo de *deep learning*, que es una rama de la inteligencia artificial y el aprendizaje automático, que utiliza redes neuronales profundas (con una gran cantidad de capas) para entrenar y procesar grandes cantidades de datos. En el caso del reto, el modelo de *deep learning* a utilizar para identificar el ventrículo izquierdo ya sea a partir de la utilización de máscaras o de landmarks, será un modelo CNN (Convolutional Neural Networks) lo cual se traduce a Redes Neuronales Convolucionales, es un tipo de red neuronal que funciona de tal forma que procesa los datos en forma de cuadrículas, algo especialmente útil cuando se trabaja con imágenes y patrones. La arquitectura de estos modelos consta de tres tipos de capas, las cuales comienzan por las capas convolucionales que se encargan de recibir la imagen y detectar características y patrones específicos de lo que se esté entrenando. Después vienen las capas de *pooling*, las cuales reducen el número de parámetros de la red (conservando las características más importantes) con el fin de reducir la complejidad del modelo y mejorar la eficiencia del modelo. Finalmente las capas completamente conectadas, las cuales ayudan en la clasificación de imágenes basada en las características detectadas anteriormente. (Awati, 2023)

Para el caso de la experimentación en el modelo de entrenamiento con *landmarks*, se probó una cierta cantidad de modelos hasta llegar al que mejores resultados arrojó, para dar un mayor contexto de los modelos utilizados, se dará una breve explicación teórica de cada modelo probado en la experimentación para obtener el mejor modelo de *deep learning* en la identificación del ventrículo izquierdo por medio de landmarks:

U-Net: Se define a este modelo como una arquitectura de CNN usada mayormente en tareas de procesamiento y segmentación de imágenes. Esta arquitectura tiene dos principales partes, la primera consta de una ruta de “contracción”, la cual realiza operaciones de convolución y pooling, la segunda parte importante se traduce en una ruta de “expansión”, donde realiza operaciones de convolución y upsampling para aumentar la resolución al mismo tiempo que reduce la profundidad de las características. (Ronneberger y cols., 2015)

VGG16: Es un modelo de redes neuronales convolucionales, que consiste en 16 capas, con 13 capas convolucionales y tres capas completamente conectadas. Este tipo de CNN es usada sobre todo por su simplicidad y en el caso del reto, se probó como una forma de aplicar un modelo diferente en el entrenamiento con landmarks. (GeeksforGeeks, 2023)

RESNET50: Este también es un modelo de CNN, donde parte de un modelo RESNET clásico el cual funciona de forma que tiene la arquitectura de una CNN tradicional con la adición de conexiones residuales”, las cuales sirven para permitir el flujo directo de información sin procesar desde las capas de entradas hasta las capas de salida. En caso de una RESNET50, tiene una profundidad de 50 capas donde se incluyen capas de entrada, convolucionales, conexiones residuales y capas completamente conectadas. Boesch (2023)

Data Augmentation: Se puede decir que data augmentation es una técnica utilizada en modelos de aprendizaje automático y aprendizaje profundo. Utilizando esta técnica, es posible reducir el sobreajuste del modelo. En el caso del reto, la manera en que funciona un data augmentation es que se aplica a las imágenes para que cambien su sentido (inversas, diagonal, etc) y de esta forma el modelo pueda aprender con diferentes ángulos de las imágenes para llegar a resultados más realistas. (Awan, 2022)

Para el caso del reto actual, es necesario aplicar un marco de trabajo para generar los modelos de *deep learning*, en este caso, se utilizó *Tensorflow*, que es una biblioteca de código abierto para implementar y desarrollar los modelos de *deep learning*, siguiendo una metodología clara para generar modelos. Para el caso del código, se utilizaron Jupyter Notebooks para la aplicación del código y visualización de resultados.

2.5. Evaluación de modelos

Es necesario recalcar que uno de los objetivos generales del reto es el de comparar la detección del ventrículo aplicando máscaras y landmarks, por lo que al realizar sus respectivos modelos de *deep learning*, debe de considerarse que la selección del mejor modelo esté condicionada por distintos aspectos, comenzando por las métricas utilizadas en la evaluación de los modelos:

Dice Score: El *Dice Score* dentro de un modelo de redes neuronales convolucionales, es la métrica utilizada para medir el rendimiento en tareas como la segmentación de imágenes y se utiliza para medir la similitud o superposición de una imagen y su máscara, por lo que la métrica en un rango de 0 a 1 (siendo

0 una indicación de no similitud y 1 de similitud perfecta), indica qué tan bien las predicciones del modelo se alinean correctamente con los datos actuales. De una forma técnica, la fórmula para obtener esta métrica sería la siguiente:

$$DICE = \frac{2 \times |X \cap Y|}{|X| + |Y|} \quad (1)$$

En la fórmula anterior, X representa el número de píxeles predichos, mientras que Y representa los píxeles originales. (Researchgate, s.f.)

En el reto, el *dice score* servirá para evaluar la manera en que el modelo predice las máscaras del ventrículo y cómo esas predicciones se comparan con las máscaras ya generadas.

Dice Loss: Esta métrica sirve como una función de pérdida que se utiliza en tareas donde se requiere segmentación de imágenes y busca medir la similitud entre dos conjuntos. Dentro del contexto del reto, se busca que al obtener y entrenar un modelo, se busque minimizar la función de pérdida ya que eso significa que las predicciones se acercan a las referencias. En caso de la fórmula de ésta función de pérdida, es la misma que la descrita en el *Dice Score* solo que vendría a ser 1 menos la fórmula del *Dice Score*.

Accuracy: Esta métrica sirve en un modelo de forma que dice qué tan bien el modelo predice comparando con los datos originales. Se usa como una métrica para saber y conocer el estado en que se encuentra un modelo para poder predecir correctamente. (Gad, 2021)

Adaloss: Es una función de pérdida adaptativa diseñada para mejorar la precisión de la localización de *landmarks* (es una función diseñada específicamente para abordar problemas relacionados con la identificación de *landmarks*), sobre todo en el ámbito de *computer vision* y tareas de *deep learning*. Esta función de pérdida se adapta durante el entrenamiento mediante la actualización de la precisión del objetivo en función del progreso del entrenamiento, lo que permite un aprendizaje con mayor efectividad. (Teixeira y cols., 2019)

Mean Squared Error: Se utiliza para medir las discrepancias entre las predicciones del modelo y las etiquetas reales de los datos de entrenamiento. La manera en que se calcula está métrica es midiendo la media de los cuadrados de las diferencias entre las predicciones y los datos reales y al elevar al cuadrado, se penalizan más los errores más grandes. Una representación sobre cómo se calcula se puede apreciar con su fórmula:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

En la fórmula, n es el número de muestra, y_i es el dato real y \hat{y}_i representa la predicción. (Gupta, 2023)

Otras funciones de pérdida y métricas probadas: Dentro de la parte de experimentación con *landmarks* se menciona que en el proceso, se realizaron diferentes pruebas con modelos, sus respectivas métricas y funciones de pérdida, tal y como se observa en la Figura 19. Por lo tanto es de valor que se pueda ofrecer una explicación resumida sobre las diferentes métricas probadas: se utilizó Binary Cross Entropy, adecuada cuando se trata de clasificar entre dos clases. Mide la discrepancia entre la distribución de probabilidad predicha y la distribución de probabilidad real para problemas de clasificación binaria. Se utilizó también Categorical Cross Entropy, mide la discrepancia entre las distribuciones de probabilidad predichas y reales

para múltiples clases. La Kullback Leiber, mide cuánta información se pierde al aproximar una distribución de probabilidad a otra, se utiliza para comparar la distribución predicha por el modelo con la distribución real de los datos. Se tuvo también Jaccard Loss, la cual se utiliza para evaluar la superposición entre la máscara de segmentación predicha y la máscara de referencia, cuanto mayor sea el índice de Jaccard, menor será la pérdida. Otra función de pérdida utilizada fue Cosine Similarity, que se puede utilizar para medir la similitud entre la representación vectorial de la predicción del modelo y la representación vectorial de los datos reales. (Pandit y Pandit, 2023)

Calidad de máscaras y *landmarks* obtenidos: Una forma para evaluar y determinar la calidad de los modelos, es mediante un análisis cualitativo de las máscaras y de los *landmarks* generados, donde se puedan comparar las predicciones y las imágenes de referencia. Con esto se puede obtener otro punto de vista que ayude a determinar si las predicciones de los modelos coinciden con los ventrículos y las imágenes de referencia.

Rendimiento de modelos utilizando bases de datos externas: Al contar con los modelos finales seleccionados para generar máscaras y *landmarks*, sería prudente y necesario probar ambos modelos con ecocardiogramas externos, es decir, con bases de datos externas para observar la calidad de predicciones hechas y ayudar a determinar qué modelo aprendió mejor.

Comparación de modelos mediante máscaras: Después de generar ambos modelos para identificar el ventrículo izquierdo, una forma de poder decir cuál fue el mejor modelo, es mediante la comparación en la generación de máscaras de ambos modelos, donde mientras el modelo de máscaras ya da máscaras como predicciones, en el modelo de *landmarks* se obtendrán los *landmarks* y después de convertirán en máscaras para compararlas con las generadas en el primer modelo. De esta forma se obtendrá un *dice score* para afirmar qué modelo fue mejor.

3. Experimentación

3.1. Modelo para Máscaras

Para comenzar con la formulación y generación del modelo de deep learning para generar máscaras del ventrículo izquierdo del corazón, se comenzó por la propia generación de las máscaras para cada frame de los videos de la base de datos original, tal y como se explica en la sección de Metodología y se observa en la Figura 7 y Figura 8. Al tener carpetas con las imágenes y máscaras generadas dependiendo si son parte del conjunto de entrenamiento, validación o prueba, se procedió a convertir cada carpeta con las imágenes o máscaras en un numpy array, debido a que este formato ofrece una forma eficiente de almacenar datos y eso tiene como consecuencia que se sea más sencilla la manipulación y extracción de las imágenes para realizar los entrenamientos de los modelos.

Al tener las imágenes y máscaras en un formato más versátil y accesible, se procedió a aplicar una binarización de las máscaras, esto para aplicarlo de tal forma que en las máscaras se tenga una región

de interés (ventrículo izquierdo) y la binarización ayude a que se identifiquen los píxeles relevantes y no relevantes en cada máscara. Para implementarlo en el modelo, se definieron funciones de coeficiente *Dice*, tanto para medir el rendimiento de las predicciones como medir la función de pérdida del modelo.

La formulación del modelo consta de agregar distintas capas. Primeramente, considerando que el tamaño de las imágenes es de 112x112 píxeles y tienen 3 canales. Con esto se definió el modelo agregando capas convolucionales, capas pooling, normalización por lotes y capas, Dropout para regularizar el modelo (Srivastava y cols., 2014) y evitar el sobreajuste. Seguido de esta serie de capas, se agregó una capa Flatten para para aplanar los datos y después aplicar capas completamente conectadas, para finalmente hacer una redimensión y obtener una salida que tenga el mismo tamaño que la entrada de las máscaras, la cuales tienen un tamaño de 112x112 píxeles con un solo canal de profundidad.

Al configurar el modelo, se agrega un punto de control (*Checkpoint*) para guardar la configuración del mejor modelo generado durante el entrenamiento, utilizando el *Dice Score* como métrica. El mejor modelo se guarda en un archivo .h5. Para la compilación del modelo se utilizó un optimizador y métricas como la función de pérdida *Dice Loss* y *Dice Score* para evaluar el rendimiento del modelo. Finalmente se comenzó con el entrenamiento del modelo, utilizando imágenes de entrenamiento y sus respectivas máscaras, se aplican 30 épocas de entrenamiento y se aplican la métricas para supervisar el modelo en el entrenamiento.

3.2. Modelo para Landmarks

Para iniciar la formulación y desarrollo del modelo de aprendizaje profundo destinado a la generación de *landmarks* en lugar de máscaras para el ventrículo izquierdo del corazón, se llevó a cabo la generación de *landmarks* para cada fotograma de los videos contenidos en la base de datos original, conforme se describe en la sección de Metodología y se ilustra en la Figura 17 y Figura 18. Una vez organizadas las imágenes y *landmarks* en carpetas específicas según su pertenencia al conjunto de entrenamiento, validación o prueba, se procedió a transformar cada carpeta conteniendo imágenes o *landmarks* en un *array* de *Numpy*. Esta elección se fundamenta en la eficiencia que este formato ofrece para el almacenamiento de datos, lo cual facilita la manipulación y extracción de imágenes necesarias para llevar a cabo los entrenamientos de los modelos.

La conversión a formato Numpy se considera esencial debido a su capacidad para representar datos multidimensionales y su eficacia en términos de velocidad y memoria. Al adoptar este enfoque, se optimiza la administración de las imágenes y landmarks, proporcionando un medio más eficiente para su procesamiento durante las fases de entrenamiento del modelo de aprendizaje profundo.

Cabe destacar que la generación de landmarks implica un paso crucial en el proceso, ya que los landmarks actúan como puntos de referencia anatómicos clave en las imágenes del ventrículo izquierdo del corazón. Estos landmarks servirán como base para el desarrollo del modelo de aprendizaje profundo, contribuyendo a la comprensión y representación precisa de la morfología cardíaca en el conjunto de datos de interés.

En el proceso de desarrollo del modelo, se reconoce la necesidad de explorar y evaluar diversas métricas para optimizar el rendimiento del sistema. Para la función de pérdida, se considerará una variedad de

opciones, incluyendo, pero no limitándose a, binary crossentropy, categorical crossentropy, mean squared error (mse), mean absolute error (mae), jaccard loss, así como funciones personalizadas de pérdida que combinan distintas métricas. La elección de la función de pérdida es fundamental para guiar al modelo hacia la convergencia óptima, y la experimentación con diferentes funciones permitirá determinar cuál se adapta mejor a la naturaleza específica del problema de generación de landmarks.

En paralelo, se explorarán diversas métricas de evaluación para garantizar una evaluación completa y precisa del rendimiento del modelo. Se considerarán métricas como la exactitud (accuracy), el índice de similitud de Dice (Dice Score) y la distancia de Jaccard, que ofrecen medidas específicas y complementarias de la calidad de las predicciones del modelo en relación con los landmarks esperados.

A continuación se estará observando un diagrama con las diferentes combinaciones de modelos, funciones de pérdida y métricas utilizadas para generar un modelo útil pero que desafortunadamente los resultados obtenidos con estas combinaciones no fueron las óptimas y no entregaron los resultados esperados.

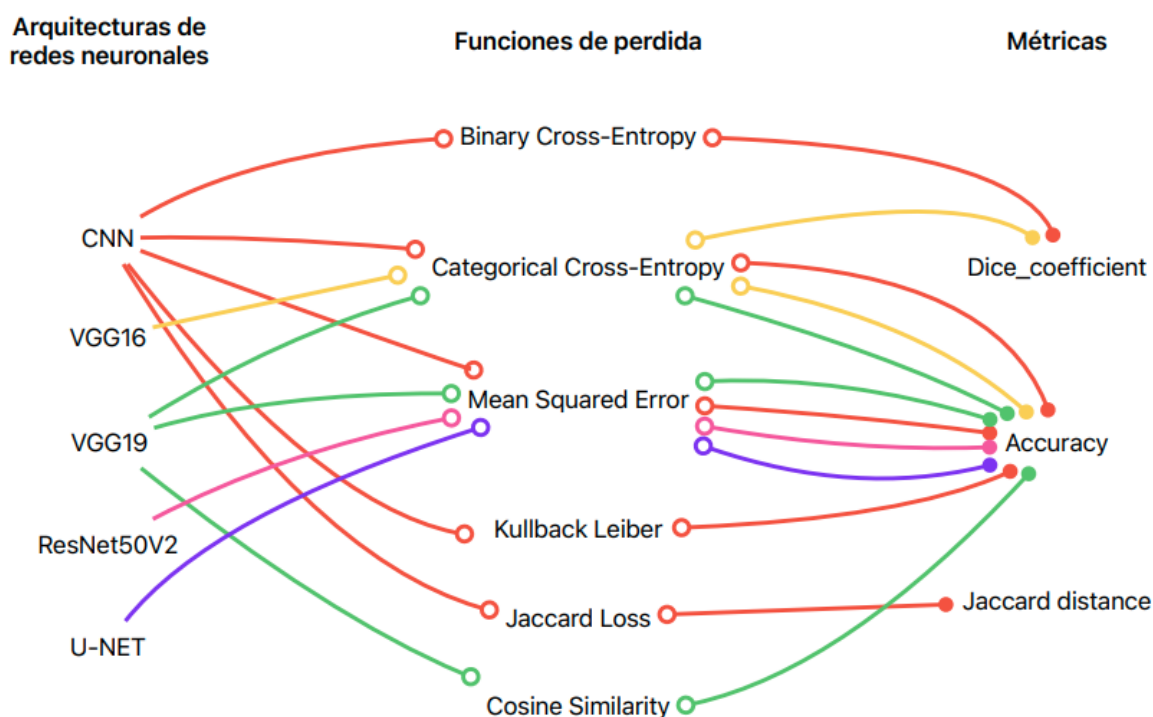


Figura 19: Modelos probados en fase de experimentación, junto con funciones de pérdida y métricas.

En el flujo de procesamiento de datos, se inició estableciendo las rutas y cargando conjuntos de entrenamiento, validación y prueba para imágenes originales y mapas de calor. Durante este paso, se aplicaron conversiones esenciales, como cambiar el tipo de datos de uint8 a float32, y se realizó una normalización

dividiendo cada valor de píxel por 255, garantizando así que los datos estuvieran en un formato óptimo para el entrenamiento de modelos de visión por computadora.

Posteriormente, se implementó una función, `replace_nan_with_zero(array)`, que desempeña un papel crucial al reemplazar los valores NaN por 0 en los mapas de calor. Este proceso es esencial para mantener la consistencia y fiabilidad de los datos, asegurando que no haya valores no válidos que podrían afectar el rendimiento del modelo.

La verificación de la presencia de valores NaN por canal en los conjuntos de datos de imágenes y mapas de calor proporcionó una evaluación detallada de la calidad y la integridad de los datos. Este análisis específico por canal permite identificar posibles problemas en datos específicos y garantizar que no haya irregularidades que puedan afectar el entrenamiento del modelo.

Se prosiguió obteniendo el tamaño de cada conjunto de datos, crucial para configurar correctamente modelos de aprendizaje automático. Los tamaños resultantes indican dimensiones consistentes en todos los conjuntos de datos de entrenamiento, validación y prueba para imágenes y mapas de calor.

Los resultados fueron los siguientes : Tamaño de `imagenes_train`: (14918, 112, 112, 3) Tamaño de `heatmaps_train`: (14918, 112, 112, 7) Tamaño de `imagenes_val`: (2576, 112, 112, 3) Tamaño de `heatmaps_val`: (2576, 112, 112, 7) Tamaño de `imagenes_test`: (2552, 112, 112, 3) Tamaño de `heatmaps_test`: (2552, 112, 112, 7)

Antes de llevar a cabo el entrenamiento del modelo, se llevaron a cabo evaluaciones adicionales identificando valores máximos y mínimos en cada conjunto de datos. La consistencia de estos resultados, donde se encontró que el valor mínimo era 0 y el valor máximo era 1 en todos los conjuntos, refleja la efectividad del proceso de preprocesamiento. Esta uniformidad en la escala de los datos es crucial para el desarrollo y la evaluación efectivos de modelos de aprendizaje automático, ya que garantiza una entrada coherente y propicia para obtener resultados fiables y significativos.

Ahora, con los conjuntos de datos debidamente preprocesados, se avanza hacia la fase de entrenamiento del modelo. Se ha implementado una clase de generador de datos, denominada `DataGenerator`, utilizando la funcionalidad proporcionada por TensorFlow a través de `Sequence`. Esta clase actúa como un generador eficiente de lotes de datos durante el entrenamiento del modelo.

La clase `DataGenerator` se ha diseñado para recibir conjuntos de imágenes y mapas de calor, permitiendo un fácil manejo de lotes de datos durante el entrenamiento. Se inicializa con los conjuntos de datos, el tamaño de lote predeterminado (16 en este caso) y se genera un arreglo de índices para facilitar el acceso a los datos de manera ordenada. La función `__len__` devuelve el número total de lotes en función del tamaño de lote establecido, y `__getitem__` proporciona un lote específico en función del índice.

Esta implementación del generador de datos es esencial durante el entrenamiento del modelo, ya que permite el flujo continuo de datos en lotes, optimizando así la eficiencia de la memoria y acelerando el proceso de entrenamiento. La adaptación de conjuntos de datos preprocesados en un formato compatible con el entrenamiento del modelo es un paso crucial para lograr resultados precisos y eficaces.

Se ha incorporado una función de callback, `AdaLossCallback`, diseñada específicamente para mejorar la adaptabilidad del modelo durante el entrenamiento. Esta clase extiende las funcionalidades proporcionadas por la biblioteca TensorFlow y juega un papel crucial en la implementación de Adaptive Loss (AdaLoss).

La clase `AdaLossCallback` incluye métodos para calcular la pérdida, tanto en términos de Media del Error Cuadrático (MSE) como para actualizar variables de adaptación en función de las variaciones en la pérdida. La pérdida se calcula en cada época y se almacena en una lista. Además, se mantiene un seguimiento de la variación de la pérdida en cada landmark.

Durante la finalización de cada época (`on_epoch_end`), se calcula la media de la pérdida acumulada y se actualiza una lista de pérdidas por landmark. Se aplica un enfoque de ventana deslizante para seguir la evolución de la pérdida. La variación de la pérdida se utiliza para actualizar dinámicamente las variables del generador de datos (`train_datagen` y `val_datagen`) con el objetivo de adaptar el modelo en tiempo real.

Este diseño refleja una estrategia de adaptación continua durante el entrenamiento del modelo, permitiendo que el sistema responda de manera dinámica a los cambios en la pérdida y, por ende, mejore la eficacia del entrenamiento del modelo en tareas específicas.

En el contexto de la preparación para el entrenamiento del modelo, la elección de arquitecturas de red neuronal juega un papel crucial. La red U-Net, conocida por su estructura de encoder-decoder, es especialmente efectiva para tareas de segmentación gracias a sus conexiones laterales, que facilitan la preservación de detalles importantes durante el proceso de encoding y decoding. La utilización de capas Dropout contribuye a evitar el sobreajuste al introducir cierta aleatoriedad en el modelo.

Por otro lado, el modelo DeepLab+ se destaca por su sofisticada incorporación de la técnica ASPP, que permite la atracción y expansión de campos receptivos para mejorar la capacidad del modelo para capturar información contextual. La utilización de la red ResNet50 como base aporta la ventaja de aprovechar representaciones de características aprendidas en conjuntos de datos masivos (ImageNet).

Ambas arquitecturas han sido configuradas con capas de Batch Normalization para acelerar la convergencia del modelo al normalizar las activaciones entre capas, y funciones de activación ReLU para introducir no linealidades. Además, la implementación de Adaptive Loss (AdaLoss) y la función de callback `AdaLossCallback` agregan una dimensión adaptativa al proceso de entrenamiento, permitiendo que el modelo ajuste dinámicamente su aprendizaje en función de la evolución de la pérdida.

En la fase de implementación del modelo, se ha seleccionado el modelo DeepLab+ con una entrada de tamaño (112, 112, 3) y 7 clases. Para optimizar el proceso de entrenamiento y mejorar la generalización del modelo, se han aplicado diversas estrategias. Entre estas estrategias se incluyen el uso de técnicas de Early Stopping para detener el entrenamiento cuando la pérdida en el conjunto de validación deja de mejorar, el guardado del mejor modelo según la pérdida en el conjunto de validación, y la reducción de la tasa de aprendizaje en caso de estancamiento.

Además, se ha implementado la función de callback `AdaLossCallback`, que introduce una adaptabilidad dinámica durante el entrenamiento. Esta adaptabilidad se basa en la variabilidad de la pérdida a lo largo del tiempo y se utiliza para ajustar los hiperparámetros del modelo de manera automática.

El proceso de entrenamiento ha comenzado con un calentamiento, utilizando un optimizador Adam con una tasa de aprendizaje de 1×10^{-4} . El entrenamiento se ha llevado a cabo durante 20 épocas, con un conjunto de datos de entrenamiento (`train_datagen`) y un conjunto de datos de validación (`val_datagen`), ambos generados mediante la clase `DataGenerator`.

Posteriormente, se ha realizado una visualización de la evolución de la pérdida a lo largo de las épocas, tanto en el conjunto de entrenamiento como en el de validación. Estos resultados proporcionan información valiosa sobre el desempeño y la convergencia del modelo.

Cabe destacar que se han incorporado callbacks adicionales, como la monitorización remota del entrenamiento a través de Dweet.io. Finalmente, se ha cargado el mejor modelo obtenido durante el entrenamiento para su posterior evaluación o uso en predicciones. Este enfoque integral y estratégico demuestra una cuidadosa preparación y seguimiento del proceso de entrenamiento del modelo DeepLab+.

4. Resultados

4.1. Modelo para Máscaras

En el modelo para mascarar se obtuvieron unos resultados bastante favorables en cuestiones de métricas, obteniendo un *Dice Score* en el conjunto de prueba de aproximadamente 0.91. No se observó alguna especie de sobreajuste en el modelo, ya que el *Dice Score* obtenido en entrenamiento y en validación se encontraba en rangos similares, a continuación se muestra el comportamiento de dichas métricas durante el entrenamiento.

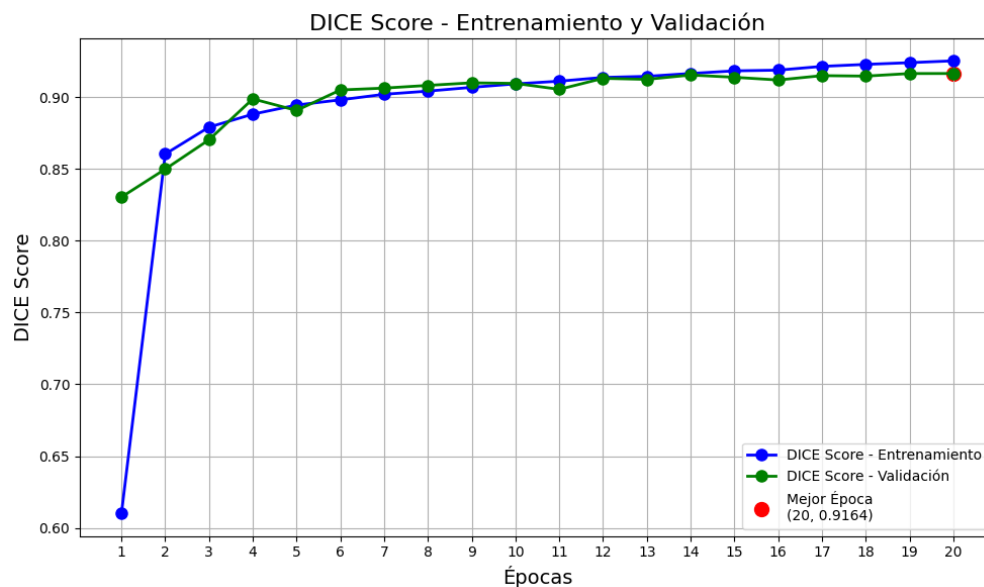


Figura 20: Dice Score durante el entrenamiento del modelo

De la misma manera, se generaron predicciones con los datos de prueba, y se comparó la máscara predicha contra la máscara real generada a partir de landmarks, con lo que se de manera visual se puede decir que se obtuvieron muy buenos resultados, pues el modelo logró adaptarse de buena manera a estos datos de prueba, como se puedes mostrar a continuación.

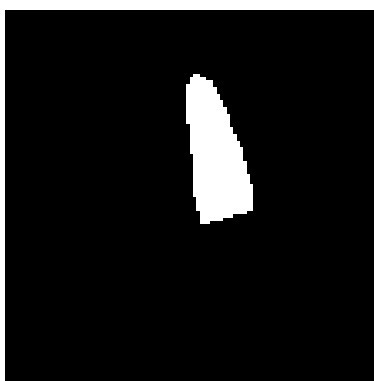


Figura 21: Mascara Original.

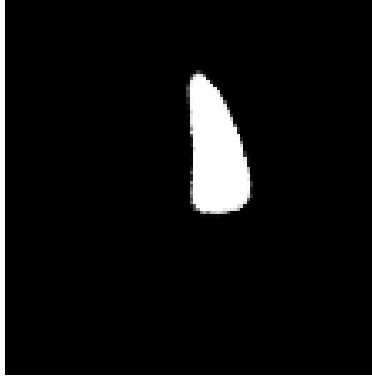


Figura 22: Mascara predicha con el modelo.

Como se puede observar en las figuras 21 y 22, las mascaras predichas son muy similares a las mascaras originales, con lo que se puede entender que el modelo es bueno, de la misma manera no se presentan sobre ajustes, pues el modelo trabaja bien tanto con los datos de entrenamieto como con los de validación, así también como los de prueba. Para comprobar que el modelo no se haya acostumbrado a los datos previamente proporcionados, se nos proporcionó una serie de videos extras de diferentes ecocardiogramas, esto con el objetivo de probar este modelo con datos nuevos, para lo cual se obtuvieron todos los frames de dichos videos y se probó, con lo que en varios casos se identificó que si identificaba de manera correcta el ventriculo izquierdo como se puede ver a continuación (Figura 23 y Figura 24).

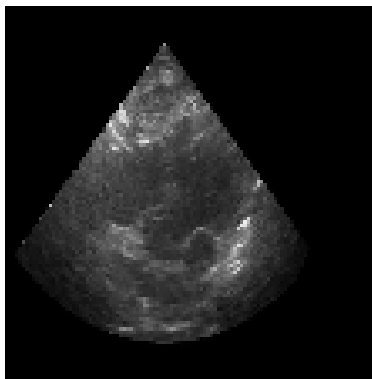


Figura 23: Imagen Original

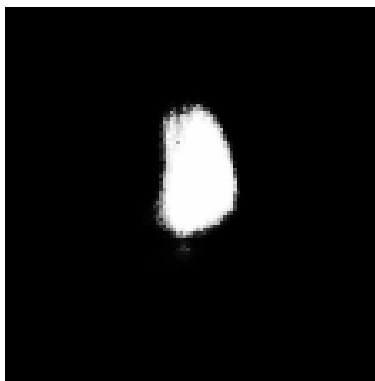


Figura 24: Mascara predicha con el modelo.

En conclusión para el análisis de resultado de este modelo, se obtuvieron resultados altamente favorables en cuanto a métricas, logrando un *Dice Score* en el conjunto de prueba de aproximadamente 0.91. La ausencia de sobreajuste se evidenció al observar que los *Dice Scores* en entrenamiento y validación se encontraban en rangos similares a lo largo del entrenamiento.

Además, las predicciones visuales también respaldaron el buen rendimiento del modelo. Al comparar las máscaras predichas con las máscaras reales generadas a partir de *landmarks*, se observó una notable similitud visual. Este resultado sugiere que el modelo se adaptó de manera efectiva a los datos de prueba, y la calidad de las predicciones se mantiene consistente.

Una comprobación adicional de la robustez del modelo se realizó al proporcionar videos adicionales de diferentes ecocardiogramas. Al extraer y probar todos los frames de estos videos, se confirmó que el modelo identifica correctamente el ventrículo izquierdo en diversas situaciones. Esta capacidad para generalizar a datos nuevos respalda la afirmación de que no hay sobreajuste en el modelo.

El modelo exhibe un rendimiento positivo en términos de métricas, calidad visual de predicciones, y generalización a nuevos datos. La consistencia en el rendimiento entre conjuntos de entrenamiento, validación, prueba y datos nuevos subraya la calidad y confiabilidad del modelo entrenado.

4.2. Modelo para *Landmarks*

A partir de pasar nuestros datos a través de la red neuronal se obtuvieron los siguientes resultados mostrados en la Figura 25, Figura 26 y Figura 27.

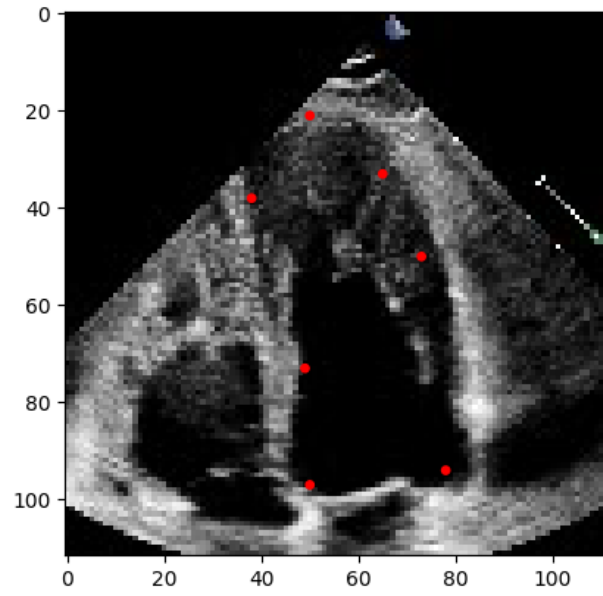


Figura 25: Landmarks predichos

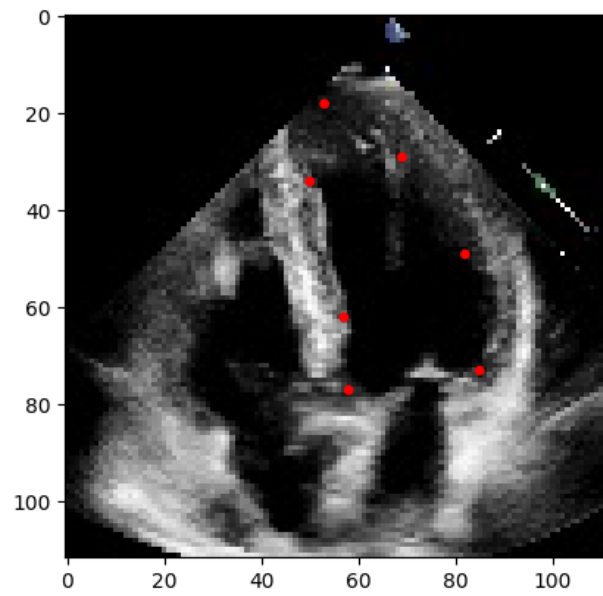


Figura 26: Landmarks predichos

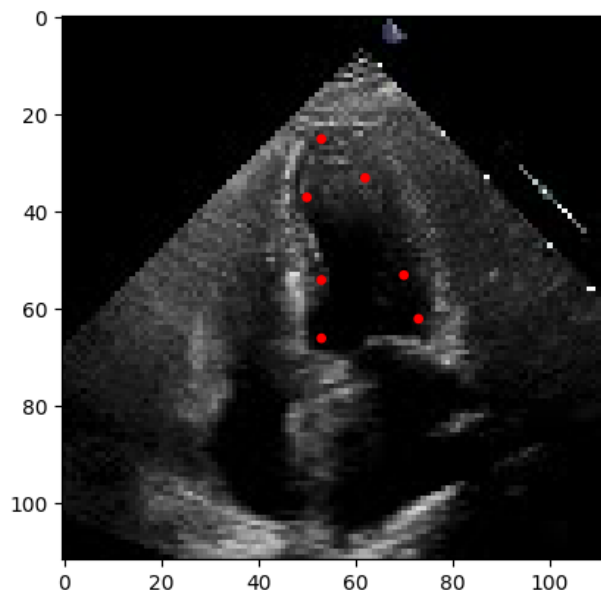


Figura 27: Landmarks predichos

Al examinar detenidamente los resultados, se observa que los landmarks predichos por la red neuronal exhiben un ajuste satisfactorio al contorno del ventrículo izquierdo en la mayoría de los casos. No obstante, se han identificado situaciones en las cuales estos puntos exceden los límites anatómicos del ventrículo izquierdo. Este hallazgo sugiere la existencia de oportunidades de mejora en el modelo de predicción.

La presencia de discrepancias señala posibles limitaciones en la capacidad del modelo para adaptarse a variaciones anatómicas menos comunes. En consecuencia, surge la necesidad de ajustar y refinar el modelo para abordar estos casos específicos y lograr una mayor consistencia en la localización de landmarks.

Este análisis objetivo destaca tanto los logros del modelo como las áreas que requieren atención y perfeccionamiento. La consideración de estas observaciones apunta hacia la implementación de un enfoque continuo de mejora, que contribuirá a fortalecer la eficacia y robustez del modelo en la identificación precisa de landmarks asociados al ventrículo izquierdo.

A continuación, se presenta el proceso detallado para obtener el `dice_score`, permitiendo así la comparación con los resultados del modelo previo, y proporcionando una evaluación cuantitativa que respalda el análisis cualitativo realizado.

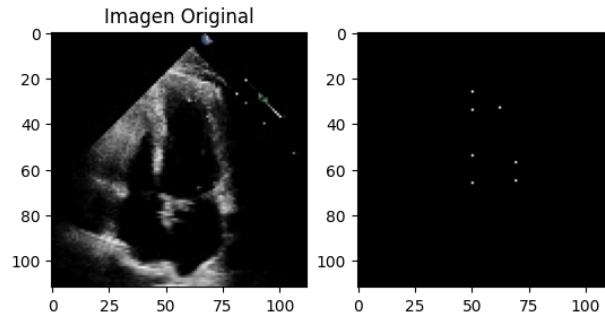


Figura 28: Imagen original y landmarks predichos

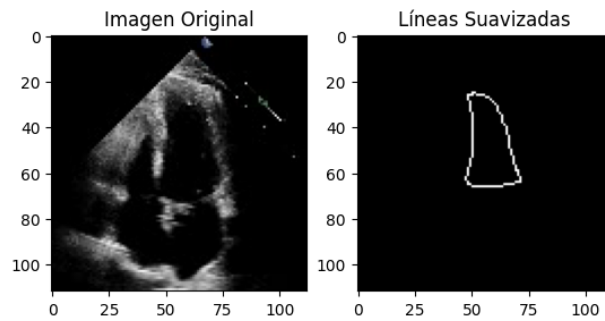


Figura 29: Imagen Original y contorno hecho a partir de landmarks

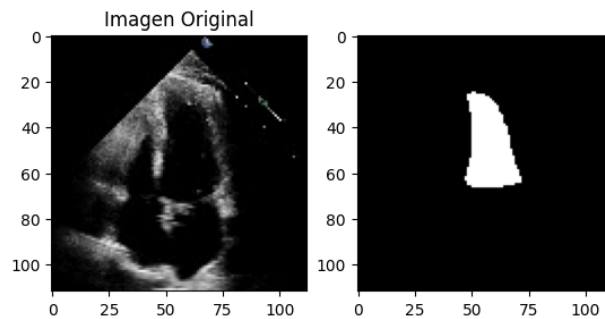


Figura 30: Imagen original y mascara final

El proceso de predicción de landmarks y el consecuente postprocesamiento para la conversión en máscaras fue llevado a cabo de manera integral, abarcando la totalidad de nuestros datos de prueba. La finalidad primordial de esta ejecución fue la obtención de un índice Dice que permitiera la comparación equitativa entre ambos modelos, utilizando una métrica uniforme.

El modelo en consideración generó un índice Dice de 0.57 como resultado de este procedimiento. Al examinar detalladamente los resultados obtenidos, se destacó que, en ciertas imágenes, el modelo lograba llevar a cabo una segmentación precisa del ventrículo izquierdo. No obstante, en otras instancias, se evidenció que el modelo no alcanzaba una segmentación correcta, presentando desafíos en la identificación y delimitación adecuada de la región de interés.

Este fenómeno observado proporciona coherencia entre el índice Dice obtenido y los resultados visuales, subrayando la variabilidad inherente al rendimiento del modelo. Se destaca que dicha variabilidad está intrínsecamente ligada a las características específicas de las imágenes de prueba, indicando la necesidad de considerar la diversidad del conjunto de datos al interpretar los resultados de la evaluación del modelo.

5. Conclusiones

En conclusión, el proyecto de segmentación del ventrículo izquierdo en ecocardiogramas mediante el uso de deep learning ha arrojado resultados altamente prometedores, especialmente en lo que respecta al modelo para máscaras.

El análisis de métricas revela un rendimiento impresionante, con un *Dice Score* en el conjunto de prueba aproximado de 0.91. La consistencia entre las métricas de entrenamiento, validación y prueba sugiere la ausencia de sobreajuste, respaldando la robustez del modelo. Las visualizaciones de las predicciones confirman la capacidad del modelo para adaptarse a datos de prueba, mostrando una similitud notoria entre las máscaras predichas y las máscaras reales generadas a partir de landmarks.

La capacidad del modelo para generalizar a nuevos datos se puso a prueba con videos adicionales, demostrando una identificación precisa del ventrículo izquierdo en diversas situaciones. Estos resultados respaldan la calidad y confiabilidad del modelo entrenado, subrayando su rendimiento positivo en términos de métricas, calidad visual de predicciones y generalización a nuevos datos.

En cuanto al modelo para *Landmarks*, se observa un ajuste satisfactorio en la mayoría de los casos, aunque con algunas limitaciones. Los landmarks predichos muestran una adaptación adecuada al contorno del ventrículo izquierdo, pero en situaciones específicas, exceden los límites anatómicos, indicando oportunidades de mejora.

La identificación de estas limitaciones destaca la necesidad de ajustar y refinar el modelo para abordar variaciones anatómicas menos comunes y lograr una mayor consistencia en la localización de landmarks. La evaluación cuantitativa mediante el índice Dice proporciona una medida objetiva que respalda las observaciones cualitativas y orienta hacia un enfoque continuo de mejora.

En resumen, el proyecto ha alcanzado éxitos notables en la segmentación del ventrículo izquierdo, estableciendo una base sólida para futuras mejoras y aplicaciones clínicas. La combinación de resultados cuantitativos y cualitativos ofrece una evaluación integral del rendimiento del modelo y señala direcciones clave para el desarrollo futuro de esta herramienta de análisis de imágenes médicas.

6. Aspectos éticos de los datos

Dentro del manejo de datos médicos, se puede considerar que existen ciertos estándares de aspectos como privacidad, acceso y uso de datos. Los estándares en general, garantizan la precisión, operabilidad y seguridad de los datos en el contexto del acceso a la información, existen diferentes estándares para el manejo de datos médicos, lo cuales pueden definirse en algunos muy relevantes como los siguientes de acuerdo con Tech (2022)

- **Health Insurance Portability and Accountability Act (HIPAA):** Se puede definir como una ley en Estados Unidos que es obligatoria para todas aquellas empresas que trabajan o manejan datos clínicos. En esta ley se establecen tres cosas, privacidad, seguridad y notificación en caso de violaciones de información o datos.

- **Health Level Seven International (HL7):** Es una organización que se encarga de desarrollar estándares para la transferencia de datos clínicos entre aplicaciones de software y proveedores de datos.

- **Digital Imaging and Communications in Medicine (DICOM):** Es un estándar internacional para el manejo e intercambio de imágenes médicas, de forma que el estándar ayuda a que las imágenes no contengan información del paciente, protegiendo su identidad.

- **General Data Protection Regulation (GDPR):** Son una serie de reglas que regulan la recolección, almacenamiento, procesamiento e intercambio de datos clínicos en las organizaciones. Las reglas buscan proteger los datos de las personas dentro de empresas.

6.1. Aplicación de estándares en los datos del reto

Los datos médicos proveídos por la base de datos EchoNet Dynamic de la Universidad de Stanford, mencionan en su página web, los términos de uso de los datos, los cuales se centran principalmente en el énfasis en la privacidad y la prohibición de distribución, publicación o reproducción de los datos a otras entidades sin permiso de la institución, esto viene a forma parte del estándar descrito en HIPAA. También se menciona una necesidad de obtención de permisos específicos para la distribución de datos, se acerca las reglas descritas en DICOM. La limitación del uso del conjunto de datos para fines de investigación no clínica y la prohibición de utilizarlo en el diagnóstico o cuidado de pacientes se alinean con los estándares éticos en la investigación médica.

Con la aplicación de diversos estándares en los datos del reto, se puede confirmar que en los datos originales dados por el socio formador, son datos que velan por la privacidad de las personas y que sobre todo son datos anónimos y que en este caso particular son datos para usarse en un proyecto académico, con la autorización y ayuda de un socio formador para su acceso, siempre manteniendo una comunicación clara para aclaración de todo tipo de preguntas relacionadas con el tema.

Referencias

- Awan, A. A. (2022, November 23). *A complete guide to data augmentation*. Descargado de <https://www.datacamp.com/tutorial/complete-guide-data-augmentation> (Accessed on November 29, 2023)
- Awati, R. (2023). *Convolutional neural network (cnn)*. <https://www.techtarget.com/searchenterpriseai/definition/convolutional-neural-network>.
- Boesch, G. (2023, November 9). *Deep residual networks (RESNet, RESNet50) – 2024 guide*. Descargado de <https://viso.ai/deep-learning/resnet-residual-neural-network/> (Accessed on November 29, 2023)
- Datalya. (s.f.). *What is landmark detection in computer vision? — data science blog*. <https://datalya.com/blog/machine-learning/what-is-landmark-detection-in-computer-vision>. (s.f.)
- Gad, A. F. (2021, April 9). *Accuracy, precision, and recall in deep learning*. <https://blog.paperspace.com/deep-learning-metrics-precision-recall-accuracy/>. (Accessed on November 29, 2023)
- GeeksforGeeks. (2023, January 10). *VGG 16 CNN model*. Descargado de <https://www.geeksforgeeks.org/vgg-16-cnn-model/> (Accessed on November 29, 2023)
- Gupta, A. (2023, February 14). *Mean squared error: Overview, examples, concepts and more*. <https://www.simplilearn.com/tutorials/statistics-tutorial/mean-squared-error>. (Accessed on November 29, 2023)
- ML-science. (s.f.). *Masking — the science of machine learning*. <https://www.ml-science.com/masking>. (s.f.)
- Now, A. (2015, 3 de 2). *Bedside ultrasound assessment of left ventricular function - acep now*. Descargado de <https://www.acepnow.com/article/bedside-ultrasound-assessment-left-ventricular-function/>
- Pandit, B., y Pandit, B. (2023, April 15). *A comprehensive guide to the 7 key loss functions in deep learning*. <https://dataaspirant.com/loss-functions-in-deep-learning/>. (Accessed on November 29, 2023)
- Researchgate, h. .fig2.341198676@ https://www.researchgate.net/figure/Dice-score-visualization-The-Dice-score-is-used-to-gauge-model-performance-ranging-from_fig2.341198676. n. . s., title = Dice score visualization. The dice score is used to gauge model. (s.f.).
- Ronneberger, O., Fischer, P., y Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. En *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, munich, germany, october 5-9, 2015, proceedings, part III* (Vol. 18, p. 234-241). Springer International Publishing.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., y Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1), 1929-1958.

- Tech, A. H. (2022, September 20). *Health data standards and their growing importance: A guide*. Descargado de <https://allhealthtech.com/health-data-standards/> (Accessed on November 29, 2023)
- Teixeira, B., Tamersoy, B., Singh, V., y Kapoor, A. (2019). Adaloss: Adaptive loss function for landmark localization. *arXiv preprint arXiv:1908.01070*.
- Tutorialsonpoint. (s.f.). *Concept of mask*. https://www.tutorialspoint.com/dip/concept_of_masks.htm. (s.f.)