



Universidad de Chile
Facultad de Ciencias Físicas y Matemáticas
Escuela de Postgrado y Educación Continua
Departamento de Ingeniería Industrial

Segmentación automática de la perfusión miocárdica en imágenes de resonancia magnética cardíaca

Tesis para optar al grado de Magíster en Ciencia de Datos
Memoria para optar al título de Ingeniero Civil Industrial

Matías Felipe Sepúlveda Godoy

Profesor guía:

Ángel Jiménez Molina

Miembros de la comisión:

Cristóbal Ramos Gómez

Sebastián Ríos Pérez

José de Grazia Kunstmann

Santiago de Chile

2023

Resumen

Las enfermedades cardiovasculares son una de las principales causas de muerte en el mundo, por lo que su detección temprana es de alto interés en la comunidad científica. En la detección de la isquemia cardíaca, una de las más relevantes dentro de las cardiovasculares, se utiliza la resonancia magnética donde es necesario segmentar los bordes endocárdicos y epicárdicos del corazón, como parte del proceso para la detección de la enfermedad, lo cuál toma un tiempo extenso.

El aprendizaje profundo es el estado del arte en la segmentación automática de la perfusión cardíaca. En este ámbito, arquitecturas como la *U-Net* adquieren gran importancia para la segmentación, y que sumadas a otras técnicas ofrecen resultados de alto nivel. En la presente investigación, se propone utilizar diferentes modelos, como la *U-Net*, *U-Net* con aprendizaje adversarial profundo y la *TransUNet*. Y de esta manera, comparar los resultados para seleccionar el modelo que permita optimizar los tiempos de ejecución con una mayor exactitud.

En el presente trabajo, se logran resultados de alto nivel en la segmentación de la perfusión cardíaca, donde se alcanzan porcentajes superiores al 90% en las métricas estudiadas, *Dice Score* e *IoU*. En particular, se destaca el 94,10% en *Dice Score* y el 94,18% en *IoU* conseguidos en el endocardio. Es importante destacar dichos porcentajes debido a que fueron obtenidos sobre la clase corregida mediante la metodología de revisión de etiquetas en imágenes de perfusión, que también es una contribución relevante de la investigación. Finalmente, se plantea profundizar en la revisión de imágenes, ya que, debido a limitaciones de tiempo y recursos, no fue posible corregir la etiqueta relacionada al epicardio, quedando en su estado previo. Lo anterior, queda pendiente como trabajo futuro, donde es necesario segmentar adecuadamente el borde del epicardio que contribuye a obtener mejores métricas en el miocardio, donde actualmente se logran resultados de un 84,08% y un 85,64% en *Dice Score* e *IoU*, respectivamente.

Agradecimientos

En la presente sección, me gustaría expresar mi agradecimiento a todas las personas que han aportado en mayor o menor medida a que hoy en día esté en posición de lograr uno de mis primeros sueños. Sin duda alguna, nada de lo que estoy por lograr en estos momentos se habría dado sin su ayuda.

Primero, me gustaría agradecer a mi madre y padre (Herminda y Luis), quienes confiaron en mis capacidades desde el principio de mis estudios. Además, sin su apoyo emocional y económico no habría tenido la oportunidad de estudiar, si hoy en día estoy en posición de alcanzar mis sueños es gracias a ustedes.

Luego, quiero agradecer a mi familia compuesta por hermanos, hermanas, cuñados, cuñadas, sobrinos y sobrinas, su cariño incondicional ha sido un soporte constante en todo este proceso, hasta el momento más pequeño donde estamos solo pasando momentos juntos ha sido un fundamental para lidiar con las frustraciones y estrés, para así poder cumplir mis sueños. De igual manera, quiero agradecer a mi pareja (Camila) su apoyo incondicional y cariño en esta etapa, que es parte fundamental para que pueda alcanzar mis sueños.

De igual manera, quiero agradecer a mis amistades de la Universidad que han sido otro pilar importante en esta bella etapa, desde mis compañeros de primer año hasta los compañeros del magíster, con los que hemos tenido muy buenos momentos tanto en los cursos de la carrera como en el tiempo libre.

Finalmente, quiero agradecer a mis profesores y profesoras, que han aportado cada uno su cariño y dedicación en la educación que he recibido en la Universidad. En especial, quiero expresar mi profundo agradecimiento a mi profesor de tesis (Ángel Jiménez), que confió en mis capacidades desde mi última práctica profesional, donde pude conocer un mundo que me apasiona como el de la investigación. Debido a esto, mis siguientes pasos serán continuar estudiando para lograr ser un gran investigador como muchos de mis profesores y profesoras.

Tabla de contenido

Introducción	1
Objetivo General	2
Objetivos Específicos	3
Estructura de la investigación	3
Marco Teórico	4
Corazón	4
Estructura	4
Enfermedades cardiovasculares	5
Imágenes médicas	6
Resonancia magnética cardiovascular	7
Cine.	8
Perfusión.	8
Análisis visual.	9
Análisis cuantitativo.	10
Aprendizaje automático	11
Aprendizaje profundo	12
Redes neuronales convolucionales.	12
Encoder-Decoder.	14
Redes generativas adversarias.	14
Transformadores.	15
Visión computacional	17
Detección de región de interés	17
Detección de saliencia.	17
Segmentación de imágenes	19
U-Net.	19
SegAN.	21
TransUNet.	22
Estado del arte	24
Redes neuronales convolucionales	24
Redes neuronales recurrentes	25
Aprendizaje adversarial profundo	25

Otros enfoques	26
Metodología	27
Revisión del etiquetado de las imágenes	27
Procesamiento previo	29
Graph-Based Visual Saliency (GBVS)	29
Modelos	30
U-Net modificada	31
U-Net modificada + aprendizaje adversarial profundo	32
TransUNet	33
Métricas	35
Coeficiente de Similitud Dice	35
Intersección sobre Unión	35
Otras técnicas	36
Transformaciones previas	36
Aumento de datos	37
Aprendizaje por transferencia	39
Regularización	39
Resultados	41
Bases de datos	41
Detalles del entrenamiento	41
U-Net	42
U-Net + Aprendizaje Adversarial Profundo	45
TransUNet	49
Resumen	51
Conclusiones	54
Contribuciones	54
Discusión	55
Trabajos futuros	56
Bibliografía	57
Anexo	61
Experimentos	61
U-Net	61

U-Net + Aprendizaje Adversarial Profundo	66
TransUNet	70

Introducción

Las enfermedades cardiovasculares son una de las principales causas de muerte en el mundo y una de las patologías que más han impactado la salud de las personas en los últimos años ([World Health Organization, 2022](#)). La Organización Mundial de la Salud (OMS), en su informe, revela que las enfermedades cardiovasculares han aumentado su mortalidad en un 25.1% entre el 2000 al 2019, causando millones de defunciones a nivel mundial ([World Health Organization, 2022](#)). La enfermedad arterial coronaria o cardiopatía isquémica es una de las principales enfermedades cardiovasculares, y su detección temprana permitiría establecer un tratamiento óptimo para el usuario.

Las organizaciones de salud suelen apoyarse en la tecnología para evaluar correctamente el funcionamiento del cuerpo humano, desde un electrocardiograma hasta un escáner. El área de cardiología cuenta con la resonancia magnética cardiovascular, que es la principal forma de diagnóstico de problemas cardíacos. Esto se debe a la capacidad de los escáneres de última generación para tomar imágenes de alta calidad con el corazón en reposo y en estrés, permitiendo capturar el borde del miocardio ([Kelle y Nagel, 2008](#)), la zona de interés para la investigación.

Las imágenes de resonancia magnética cardiovascular se dividen en dos tipos. Por un lado, se tienen las imágenes de cine que se realizan con retención respiratoria durante algunos segundos para poder ver adecuadamente el miocardio. Y por otro, se tienen las imágenes de perfusión que se hacen con la inyección de un agente de contraste mientras se le permite respirar normalmente al usuario. De esta manera, se puede diferenciar adecuadamente el miocardio del resto de la imagen, pero con un factor de complejidad adicional que es el movimiento del corazón. Los dos tipos de examen son altamente utilizados en centros hospitalarios para el diagnóstico de la enfermedad arterial coronaria.

En la práctica clínica habitual para evaluar la presencia de enfermedades cardiovasculares, se utiliza un procedimiento que implica el análisis visual de las imágenes.

Pero, la exactitud del examen depende de la experiencia del evaluador y, por lo tanto su uso es menos recomendable si es comparado con el análisis cuantitativo de los exámenes (Bremerich et al., 2013). Sin embargo, en el cuantitativo existe un proceso intermedio en donde se segmenta el miocardio, para el que se tiene que dedicar un tiempo extenso y significa un “cuello de botella” en el proceso de diagnóstico de las afecciones cardiovasculares.

Diversos algoritmos y set de datos han sido publicados en revistas de investigación, debido al creciente interés en la disciplina de inteligencia aplicada a las imágenes médicas. En particular, se han propuesto variados algoritmos de segmentación automática del miocardio, no obstante, muchos estudios se han centrado en la tarea aplicada sobre imágenes de cine y se ha prestado bastante menos atención a la tarea enfocada en imágenes de perfusión. El principal motivo de que existan menores avances es la complejidad de automatizar la tarea debido al movimiento del corazón en perfusión cardíaca, los limitados set de datos disponibles para la comunidad científica y los algoritmos con mejores resultados en la tarea general de segmentación semántica no han sido aplicados a la tarea específica en perfusión cardíaca.

El Instituto Sistemas Complejos de Ingeniería en conjunto con el Hospital Clínico de la Universidad de Chile, tienen una colaboración para desarrollar proyectos de interés mutuo, donde se puedan aplicar elementos de la ingeniería al área de salud. Uno de los proyectos consistió en la recopilación de datos de usuarios y sus respectivas secuencias de resonancia magnética cardiovascular de la perfusión miocárdica, para iniciar con el proceso de etiquetado de las imágenes y así generar algoritmos de segmentación semántica. Por lo tanto, el presente trabajo de investigación, que es la continuación de una serie de trabajos anteriores sobre los datos recopilados, tiene el propósito de aplicar técnicas más avanzadas para obtener mejores resultados en la segmentación automática de la perfusión miocárdica.

Objetivo General

El objetivo general de la presente tesis de investigación es el siguiente:

Desarrollar un algoritmo de aprendizaje profundo para la segmentación automática de la perfusión miocárdica en imágenes de resonancia magnética, que contribuya a disminuir los tiempos en el proceso de segmentación de la zona de interés.

Objetivos Específicos

Los objetivos específicos para el presente trabajo de investigación son los siguientes:

1. Implementar una arquitectura que logre mejores resultados sobre los datos de perfusión cardíaca en la tarea de segmentación semántica.
2. Comparar el desempeño de diferentes arquitecturas sobre los datos de perfusión cardíaca en la tarea de segmentación semántica.
3. Comparar el desempeño de diferentes arquitecturas sobre una base de datos externa para medir la capacidad de generalización de los algoritmos.

Estructura de la investigación

El presente informe de investigación se divide en seis secciones. En el primer capítulo de introducción, se puede comprender a grandes rasgos la importancia del problema general y el nicho al que se apunta para contribuir a mejorar las condiciones sobre el mismo. En el siguiente capítulo, que es el marco teórico, se procede a explicar los conceptos importantes para la adecuada comprensión de los siguientes capítulos. En el tercero, se presenta y analiza el estado del arte en la tarea de segmentación semántica sobre datos de perfusión cardíaca de la presente investigación. En el cuarto, se muestra la metodología planteada para resolver los objetivos mencionados anteriormente. En el penúltimo capítulo, se pueden encontrar los diferentes resultados obtenidos a través de la implementación de la metodología propuesta. Finalmente, se dejan las referencias utilizadas para construir el estudio.

Marco Teórico

En el presente capítulo, se exploran los conceptos importantes para comprender adecuadamente las siguientes secciones de la investigación. En específico, se procede a analizar desde los conceptos relacionados a la medicina, hasta las ideas técnicas en el campo de la inteligencia artificial y cómo se relaciona con las tareas médicas. De esta manera, se espera que el lector se pueda informar sobre los conceptos imprescindibles para comprender los resultados obtenidos y la contribución que genera al área específica.

Corazón

Estructura

El músculo cardíaco es un órgano imprescindible para el ser humano, por lo que su debido cuidado es una de las actividades con mayor atención en la medicina moderna. En particular, es necesario enfocarse en el miocardio, endocardio y epicardio, partes que rodean al ventrículo izquierdo como se puede ver en la Figura 1. Primero, el miocardio es la parte que bombea la sangre al cuerpo mediante su contracción, por lo que su correcto funcionamiento es vital para que no se genere alguna enfermedad. El epicardio es una capa que se adhiere al miocardio y la más externa del corazón. Y el endocardio, es la capa más interna del corazón que se encuentra junto al miocardio. En conjunto, son objeto de análisis para evaluar la presencia de alteraciones cardiovasculares como la enfermedad arterial coronaria.

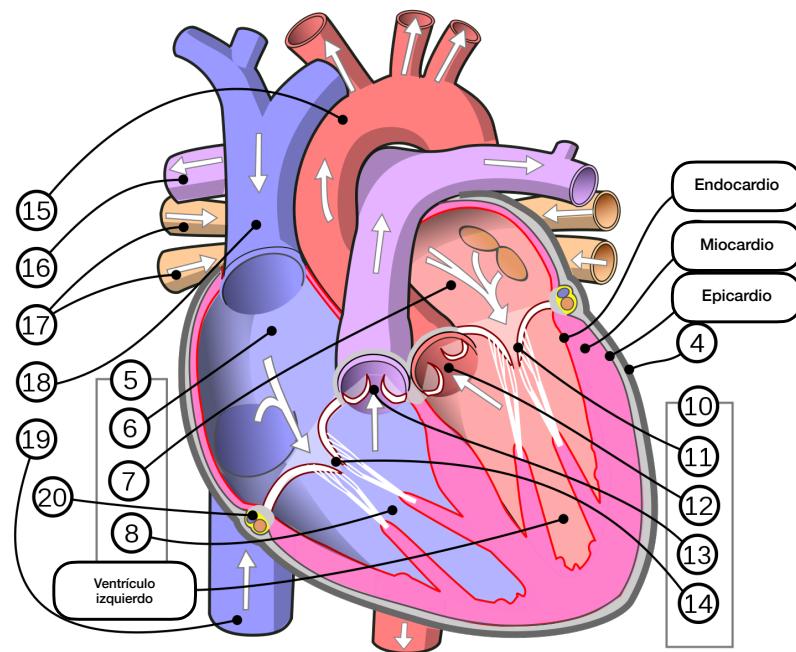


Figura 1

Estructura y Componentes del Corazón. 5. Cámaras cardíacas, 6. Atrio derecho, 7. Atrio izquierdo, 8. Ventriculo derecho, 10. Válvulas cardíacas, 11. Válvula mitral, 12. Válvula aórtica, 13. Válvula pulmonar, 14. Válvula tricúspide, 15. Arteria aorta, 16. Arteria pulmonar, 17. Venas pulmonares, 18. Vena cava superior, 19. Vena cava inferior.

Enfermedades cardiovasculares

Las enfermedades cardiovasculares son de las más relevantes a nivel mundial, especialmente en el occidente, y dentro de ellas, la principal es la enfermedad arterial coronaria o cardiopatía isquémica, la cual es parte importante de la presente investigación. En particular, se tiene que es causa de una gran frecuencia de muertes a nivel mundial (World Health Organization, 2022), por lo que su diagnóstico temprano es clave para la radiología. La

cardiopatía isquémica o enfermedad arterial coronaria es una enfermedad que ocurre cuando las arterias que suministran la sangre al músculo cardíaco se endurecen y se estrechan. Esto se debe a la acumulación de colesterol y otros materiales en la capa interna de las paredes de la arteria. Como consecuencia, el músculo cardíaco no puede recibir la sangre o el oxígeno que necesita. Con el tiempo, la enfermedad puede debilitarlo y contribuir a la presencia de insuficiencia cardíaca y arritmias ([Texas Heart Institute, 2022](#)).

Imágenes médicas

Las imágenes médicas están relacionadas a diferentes tecnologías utilizadas para ver el cuerpo humano, para poder diagnosticar, controlar y tratar un posible problema del paciente. Puede abarcar desde observar una enfermedad en el órgano estudiado o hasta ser utilizado para diseñar un tratamiento en base al problema detectado. Este tipo de imágenes es una herramienta muy importante para la medicina y existen diferentes tipos de modalidades que han ido progresando durante los años, como lo puede ser la radiografía, endoscopia, tomografía computada (CT), imágenes de mamografía (MG), imágenes de ultrasonido, imágenes de resonancia magnética (MRI), tomografía de emisión de positrones (PET), etc.

En los últimos años, la disciplina de imágenes médicas ha capturado la atención de la comunidad científica de aprendizaje profundo, debido a la complejidad de las tareas que existen y el enorme beneficio que se puede obtener mediante el uso de la tecnología aplicada. La investigación aplicada a la medicina tiene diferentes tareas, que van desde actividades como análisis de reportes clínicos masivos hasta otras más complejas como segmentar objetos en una imagen. En la Figura 2 se pueden observar algunas de las actividades más comunes del aprendizaje profundo aplicado a la salud.

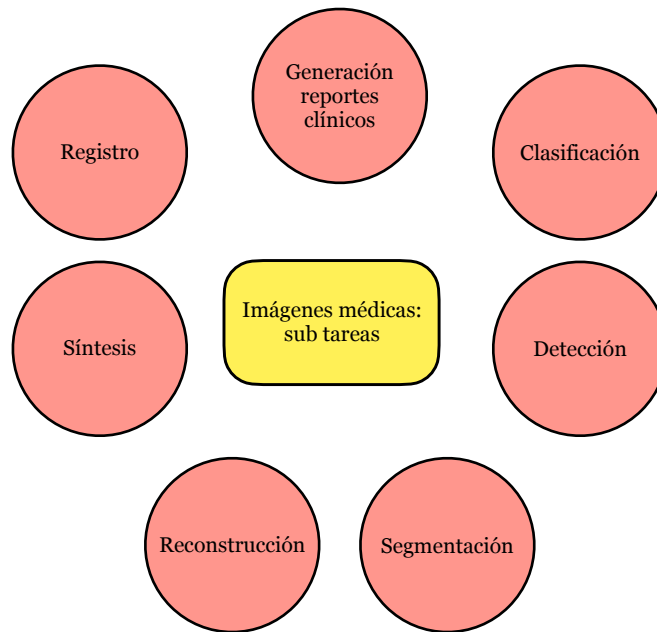


Figura 2

Diagrama de sub tareas de visión computacional y aprendizaje profundo en imágenes médicas.

Resonancia magnética cardiovascular

La resonancia magnética crea imágenes utilizando ondas de radio en presencia de un campo magnético (Kelle y Nagel, 2008). Estas son empleadas para evaluar la presencia de enfermedades cardiovasculares como la cardiopatía isquémica y tienen una gran importancia en el trabajo de cardiología, debido a su versatilidad para una evaluación fiable, amplia y profunda del funcionamiento o la estructura del corazón (Kelle y Nagel, 2008). A continuación, se explican los dos tipos de exámenes de resonancia magnética cardiovascular que son relevantes para el presente estudio.

Cine.

Las imágenes de cine son un tipo particular de resonancia magnética cardiovascular, en el que el examen se toma mientras el paciente realiza retención respiratoria durante 4-6 segundos aproximadamente. De esta manera, se puede observar el miocardio de forma estable y con los bordes del endocardio más visibles. Uno de los set de datos más conocidos de imágenes de Cine MRI, es el correspondiente a la competencia “ACDC” ([Bernard et al., 2018](#)). En este no se tiene una secuencia como tal debido a que son imágenes 3D, pero se seleccionaron dos fotogramas que si son secuenciales y que corresponden a las instancias de la fase diastólica final (izquierda) y sistólica final (derecha). A continuación, en la Figura 3 se muestra un ejemplo de las dos fases presentes en las imágenes de cine cardíaco, a la izquierda la fase diastólica y a la derecha la sistólica final.

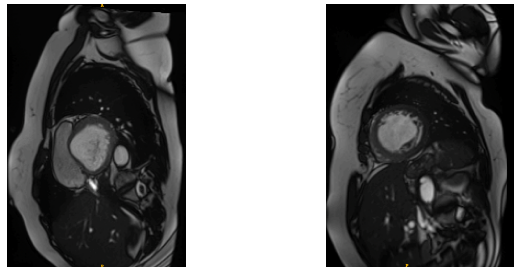


Figura 3
Secuencia de imágenes de cine.

Perfusión.

Las imágenes de perfusión son otro tipo particular de resonancia magnética cardiovascular, en el que el examen se toma mientras el paciente está respirando, un factor que aumenta la dificultad para observar adecuadamente el miocardio. Este proceso se realiza tanto en reposo como en estrés. Por ello, para realizar el examen se coloca una inyección de un agente de contraste que contiene gadolinio estándar y luego de 30-50 segundos se completan las

imágenes. En la Figura 4, se puede observar cómo va cambiando la imagen con la inyección del agente de contraste, hasta que se puede ver claramente los bordes del endocardio.

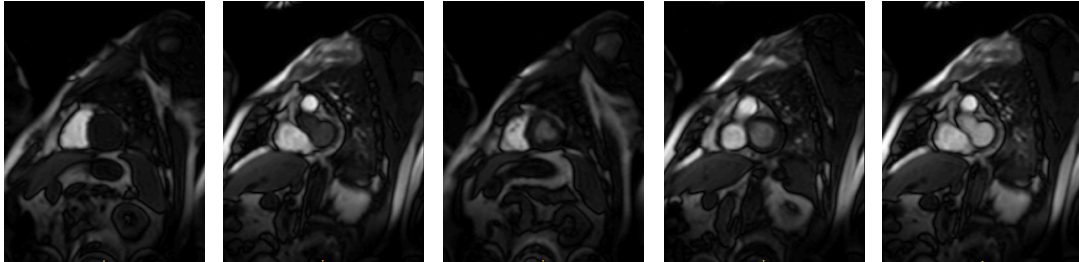


Figura 4
Secuencia de imágenes de perfusión.

Para evaluar la presencia de alguna enfermedad cardiovascular, las imágenes segmentadas de perfusión cardíaca se someten al siguiente procedimiento y análisis basado en un artículo de normas estándar ([Bremerich et al., 2013](#)).

Análisis visual.

Se enfoca en observar el defecto de perfusión en las imágenes. Para ello, se deben aplicar los siguientes lineamientos en una fase previa a la detección de una patología.

1. Visualizar las imágenes de perfusión cardíaca tomadas tanto en estrés como en reposo, una al lado de otra.
2. Ajustar la ventana y el nivel, lo que equivale a revisar que el miocardio antes de la llegada del agente de contraste, y verificar que es casi negro y que la señal capturada en el ventrículo izquierdo y derecho es de color gris brillante en lugar de blanco puro.
3. Revisar las series desplazándose de forma individual por las imágenes, considerando la ventana desde que llega el agente de contraste al miocardio hasta cuando no es visible.

Luego, para observar el defecto de perfusión en las imágenes los médicos radiólogos deben tener en consideración los siguientes aspectos.

1. Para identificar la hipoperfusión relativa, se debe comparar entre regiones endocárdicas y epicárdicas, entre segmentos del mismo corte o entre cortes.
2. Se deben contrastar las imágenes de estrés y reposo para identificar los defectos de perfusión inducibles y los artefactos.
3. Un defecto de perfusión inducible se produce cuando el contraste llega al miocardio del ventrículo izquierdo y persiste más allá del pico de realce miocárdico durante varios intervalos. Además, tiene que ser de un tamaño mayor a un pixel de ancho y es más prominente en la porción subendocárdica del miocardio.
4. Es importante notar que el defecto de perfusión inducible se manifiesta como un gradiente transmural a través del grosor de la pared del segmento afectado, es decir, más denso en el endocardio y gradualmente menos denso hacia el epicardio. Con el tiempo, el defecto retrocede hacia el subendocardio.
5. El defecto se encuentra en imágenes de estrés y no en reposo, y se ajusta al territorio de distribución de una o más arterias coronarias.

Análisis cuantitativo.

Se enfoca en complementar el análisis visual descrito anteriormente, ya que se puede realizar una descripción objetiva del campo de la intensidad de la señal en los estudios de resonancia magnética cardiovascular de perfusión miocárdica (Bremerich et al., 2013). Para poder evaluar correctamente se propone en (Bremerich et al., 2013) la siguiente metodología de análisis cuantitativo o semi cuantitativo.

1. Seleccionar una imagen de la serie dinámica con buen contraste entre todos los compartimentos cardíacos.

2. Delinear los contornos endocárdicos y epicárdicos del ventrículo izquierdo.
3. Propagar los contornos a todas las demás imágenes dinámicas.
4. Utilizar alguno o varios de los métodos que se han propuesto para el análisis cuantitativo.

En particular, se pueden observar los más relevantes en la revisión de Jerosch-Herold (2010) sobre la cuantificación de la perfusión cardíaca.

Aprendizaje automático

El aprendizaje automático es un campo multidisciplinario que contempla la utilización de computadores y algoritmos para extraer conocimiento de un determinado set de datos, y así utilizar esa información de valor para tomar decisiones. Es una disciplina que lleva muchos años en la teoría, pero que ha tenido su mayor popularidad en el último tiempo debido a la capacidad computacional que ha ido en aumento, permitiendo a los diferentes investigadores resolver problemas que antes eran muy complicados.

A continuación, se describen los tres tipos de aprendizaje automático que se han construido hasta el momento:

1. Aprendizaje supervisado: se centra en aprender las tareas para las que se tiene información histórica, e intenta predecir comportamientos nuevos en base a ese conocimiento a priori.
2. Aprendizaje no supervisado: se enfoca en la extracción de datos sin conocimiento a priori, por lo que se intenta comprender cómo es la estructura de la información y de esa manera poder ser utilizado con distintos propósitos cualitativos o cuantitativos.
3. Aprendizaje reforzado: se utiliza para aprender a realizar tareas a partir de su interacción con un determinado entorno.

Aprendizaje profundo

El aprendizaje profundo es un subconjunto del aprendizaje automático, que utilizando algoritmos de redes neuronales artificiales ha permitido importantes avances en los resultados actuales que se han obtenido en el campo de la inteligencia artificial. Las redes neuronales artificiales se basan en la idea de cómo interactúan las redes neuronales biológicas presentes en nuestro cerebro, aunque autores actuales piensan que son más una forma elegante e intuitiva de representar algo tan complejo como la interacción entre grandes cantidades de funciones flexibles, donde a cada una de estas se les denomina neuronas por simplicidad ([Hanin, 2022](#)). En términos prácticos, lo que las redes neuronales artificiales hacen es tomar los datos del mundo real y entender cómo interactúan para resolver una o varias tareas diferentes de las que se tiene conocimiento a priori o no.

En el aprendizaje profundo existen diferentes tipos de redes neuronales y estructuras de ellas, las más usuales son el perceptrón, el perceptrón multicapa, las redes neuronales convolucionales, las redes neuronales recurrentes, los transformadores, etc. En particular, nos enfocaremos en las redes neuronales convolucionales, las estructuras encoder-decoder, la estructura generativa adversaria y los transformadores.

Redes neuronales convolucionales.

Las redes neuronales convolucionales son un tipo de redes neuronales artificiales enfocadas en la extracción de características de imágenes o videos. Para ello, se utiliza una operación matemática llamada convolución, que se define como un operador que transforma dos funciones f y g en una tercera función que está relacionada a la magnitud de superposición de las funciones f y una versión trasladada e invertida de g . Esto se utiliza en las redes neuronales convolucionales para extraer diferentes patrones dentro de las imágenes que permitan a las capas finales decidir sobre la tarea objetivo. Los patrones que se pueden extraer

dependen del filtro que se utilice, que van desde un kernel para detectar líneas verticales u horizontales, hasta para detectar líneas en otros grados de inclinación menos convencionales.

En el proceso de extracción de características de una imagen, hay diferentes factores involucrados que afectan el tamaño de salida de la capa convolucional. Algunos de los más relevantes son la zancada con la que se mueve el kernel superpuesto sobre la imagen de entrada, el relleno que involucra agregar pixeles exteriores en la imagen de entrada y el tamaño del kernel que corresponde al tamaño del filtro o los filtros a utilizar. Los elementos anteriores se reúnen en la Ecuación 1 para determinar el tamaño de salida, considerando n el tamaño de entrada, f el tamaño del kernel, p el relleno y s la zancada.

$$(1) \quad \frac{n - f + 2 * p}{s} + 1$$

A continuación, se puede ver el resultado de aplicar una convolución sobre una matriz de 4x4, utilizando un filtro kernel de 2x2 que sirve para detectar los bordes verticales en una imagen, una zancada de valor 1 y sin relleno en la imagen. De esta manera, se obtienen solo valores en la columna central, lo que indica dónde está el borde vertical detectado en la imagen.

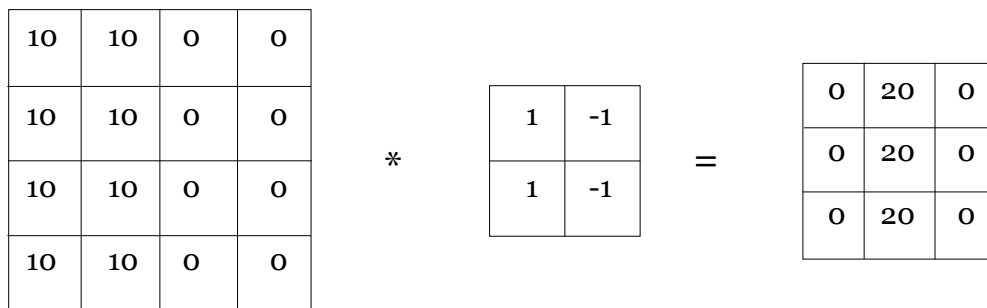


Figura 5
Ejemplo de convolución aplicada con un filtro kernel que detecta bordes verticales.

Encoder-Decoder.

Una estructura *encoder-decoder* es un tipo particular de red neuronal artificial que tiene dos componentes principales. El primer componente es un codificador, que se encarga de procesar los datos de entrada hasta obtener un espacio de variables latentes que van a ser la entrada del *decoder*. El segundo es un decodificador, que se encarga de producir una secuencia de longitud variable a partir de lo codificado anteriormente. La estructura doble es ampliamente utilizada en tareas donde se quiere construir una respuesta con una estructura similar a la de la entrada, como generar texto en tareas de lenguaje natural o imágenes en visión computacional.

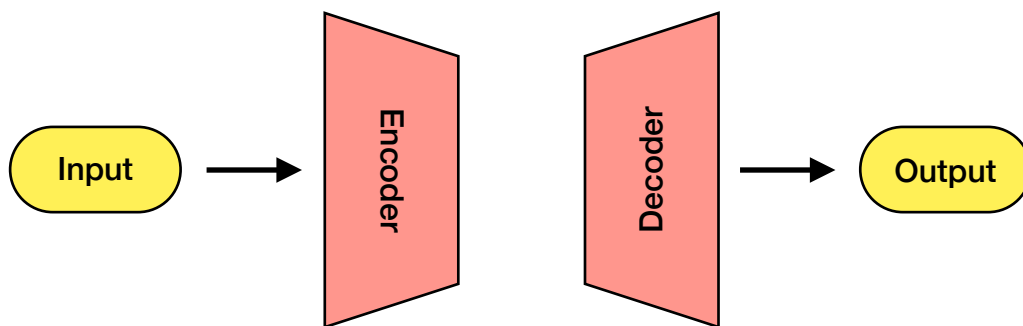


Figura 6
Estructura encoder-decoder.

Redes generativas adversarias.

Es un tipo de arquitectura particular diseñada originalmente para la generación de imágenes que parecen reales ([Goodfellow et al., 2014](#)) y que luego se fue expandiendo a otros campos como la música o segmentación semántica. Se compone de dos arquitecturas, una red neuronal generadora y otra que discrimina si lo que le llega como entrada es real o falso. Además, se denominan redes adversarias porque ambas redes son entrenadas con un propósito

distinto. La red generativa tiene el objetivo de engañar a la red discriminadora maximizando la probabilidad de que cometa un error y, por otro lado, la red discriminadora tiene el objetivo de categorizar de mejor manera si una muestra es real o falsa. En la Figura 7, se puede observar un esquema general de lo que es una red generativa adversaria desde la entrada, que es un vector de valores (en ocasiones llamado ruido), hasta la salida del discriminador donde se clasifica si el ejemplo ingresado es verdadero o falso.

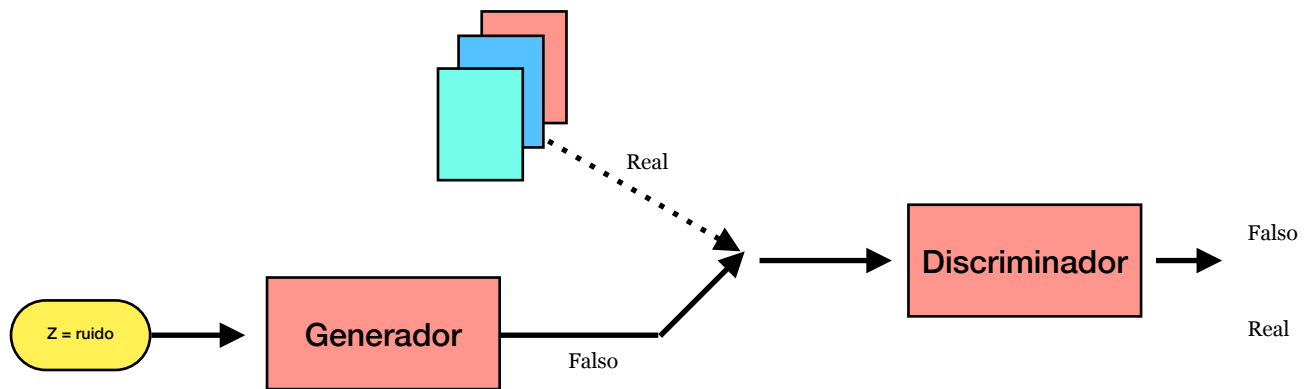


Figura 7
Red generativa adversaria.

Transformadores.

Es un tipo de arquitectura diseñada para trabajar con datos temporales, que originalmente se planteó como un método de traducción automática, pero que se propagó rápidamente a diferentes tareas (Vaswani et al., 2017). Los modelos de transformadores se basan en el concepto de atención, que se define como la asignación de una consulta y un conjunto de pares clave-valor a una salida, donde cada uno de estos elementos son vectores. La salida es una suma ponderada de los valores, donde el peso asignado a cada uno se calcula mediante una función de compatibilidad de la consulta con la clave correspondiente. De esta manera, la atención ayuda a procesar numéricamente cómo se relacionan los diferentes

componentes de la secuencia de entrada. Además, en el artículo se propone la arquitectura de la Figura 8 que consiste en un *encoder-decoder*. Este introduce la atención multicabeza, correspondiente a varias capas de atención procesadas en paralelo que posteriormente son concatenadas.

Previo a los transformadores, las redes recurrentes y convolucionales fueron el estado del arte en las tareas secuenciales, sin embargo presentaban dos problemas importantes. Primero, el costo computacional era alto, debido a que no se podía realizar el procesamiento en paralelo de las secuencias de un entrenamiento. Segundo, una limitada memoria de largo plazo, ya que les costaba comprender relaciones entre componentes de secuencias extensas al tener el limitante de realizar el cálculo de manera sucesiva y la capacidad acotada de memoria en los lotes.

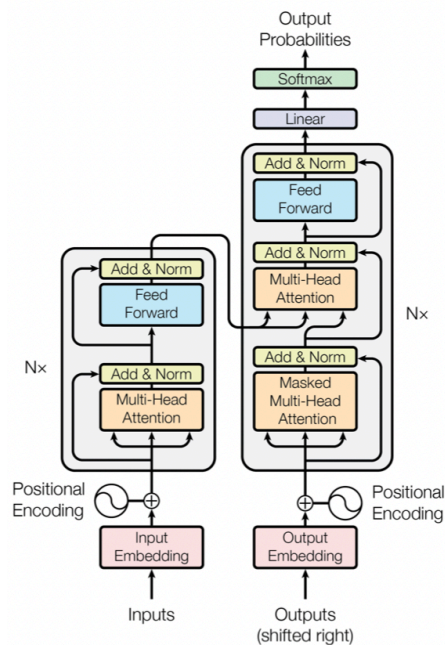


Figura 8
El transformador - modelo de arquitectura.

Visión computacional

La visión computacional es un campo de la inteligencia artificial enfocada en la extracción de información significativa de imágenes o videos. De esta manera, se pueden tomar decisiones a partir de la información extraída, su uso es muy amplio y va desde clasificar imágenes en una determinada clase hasta sistemas complejos de extracción de información del entorno como en conducción autónoma. En visión computacional, existe una gran variedad de tareas, desde las tareas tradicionales como clasificación, detección o segmentación hasta las más nuevas como seguimiento de objetos o generación de imágenes. A continuación, se procede a profundizar en las tareas relacionadas a la presente investigación, como detectar la región de interés y segmentar imágenes.

Detección de región de interés

El concepto de región de interés es ampliamente utilizado en imágenes médicas y su objetivo varía dependiendo del tipo de set de datos que se tenga. En el caso de datos en una dimensión, la tarea consiste en detectar un intervalo en la serie temporal que sea de interés. Por otro lado, con datos en dos o tres dimensiones, consiste en detectar los límites de un objeto en la imagen. Respecto a datos con tres dimensiones se deben considerar los límites asociados a profundidad. En particular, la presente investigación se enfoca en set de datos con dos dimensiones, donde se profundiza en un tipo de algoritmo específico de saliencia.

Detección de saliencia.

Los algoritmos de saliencia están relacionados a la manera en que el ojo humano detecta las partes más relevantes de una imagen, disminuyendo la atención sobre el resto de la misma. En particular, una variedad de algoritmos han sido publicados con ese objetivo de “imitar” cómo las personas tienen la capacidad de detectar las zonas salientes o regiones de interés de una

imagen. A continuación, se va a profundizar en el algoritmo *Graph-Based Visual Saliency* que es planteado en la metodología.

El algoritmo Graph-Based Visual Saliency o GBVS ([Harel et al., 2006](#)), es relativamente antiguo, pese a ello, en algunos casos sigue siendo efectivo. Así se demuestra en el artículo ([Akbari et al., 2018](#)), donde se combina con diferentes partes de procesamiento de imágenes para obtener resultados en la detección de una región de interés. A continuación, se procede a explicar el algoritmo original planteado, que se va a utilizar como parte principal en el proceso de detección de una región de interés planteado en la sección de Metodología.

El algoritmo GBVS consiste en dos fases markovianas, que son la activación y la normalización.

- Activación: se computan los mapas de activación que exponen las regiones de disimilitud conectadas. Para ello, se propone la siguiente definición de disimilitud en la Ecuación 2.

$$(2) \quad d((i, j) || (p, q)) = \left| \log \frac{M(i, j)}{M(p, q)} \right|$$

Por otro lado, se considera el grafo dirigido completamente conectado G_A , que se obtiene al relacionar cada nodo de la red M , etiquetado con dos índices $(i, j) \in [n]^2$, con todos los demás nodos $n - 1$. A cada arista dirigida del nodo (i, j) al nodo (p, q) se le asignará un peso como se indica en la Ecuación 3.

$$(3) \quad w_1((i, j), (p, q)) = d((i, j) || (p, q)) * F(i - p, j - q), \text{ donde}$$

$$(4) \quad F(a, b) = \exp\left(-\frac{a^2 + b^2}{2 * \sigma^2}\right), \text{ con } \sigma \text{ un parámetro libre.}$$

- Normalización: se concentra la masa de densidad de los mapas de activación en los lugares clave. En esta segunda fase se recibe un mapa de activación $A : [n]^2 \rightarrow \mathbb{R}$, que es lo que se desea normalizar. Se construye un grafo G_N con n^2 nodos etiquetados con índices de $[n]^2$. Para ello, cada nodo (i, j) y cada nodo (p, q) con el que se conecta, incluyendo (i, j) , se introduce una arista desde (i, j) a (p, q) con el peso indicado en la Ecuación 5.

$$(5) \quad w_2((i, j), (p, q)) = A(p, q) * F(i - p, j - q)$$

Segmentación de imágenes

La tarea de segmentación de imágenes ha capturado bastante interés en los últimos años debido a la cantidad de aplicaciones posibles en sistemas que se benefician de la automatización. Esto va desde la conducción autónoma hasta la aplicación en imágenes médicas, que son actividades que requieren de la ubicación de elementos de manera precisa en las imágenes. A continuación, se realiza un análisis más profundo de los principales algoritmos enfocados a la tarea de segmentación en imágenes médicas.

U-Net.

La arquitectura U-Net fue presentada como una red neuronal convolucional para la segmentación de imágenes biomédicas, que rápidamente se hizo relevante como el estándar de la segmentación de imágenes ([Ronneberger et al., 2015](#)). Se compone de una estructura *encoder-decoder*, donde el *encoder* se construye a partir de bloques de dos capas convolucionales con función de activación ReLU y una operación de *max-pooling*. Por otro lado, el *decoder* tiene la misma estructura, pero con una capa de convolución transpuesta para hacer el *up-sampling* en lugar de la operación de *max-pooling* que se hacía anteriormente. Además,

cada bloque del *encoder* tiene un salto de conexión hacia el correspondiente bloque en el *decoder*, para conservar características intermedias que se van obteniendo y, contribuir a una segmentación adecuada en base a ello. Finalmente, la estructura termina con una convolución de 1 x 1, para obtener la máscara de segmentación. En la Figura 9, se muestra la estructura inicialmente propuesta para la segmentar imágenes médicas.

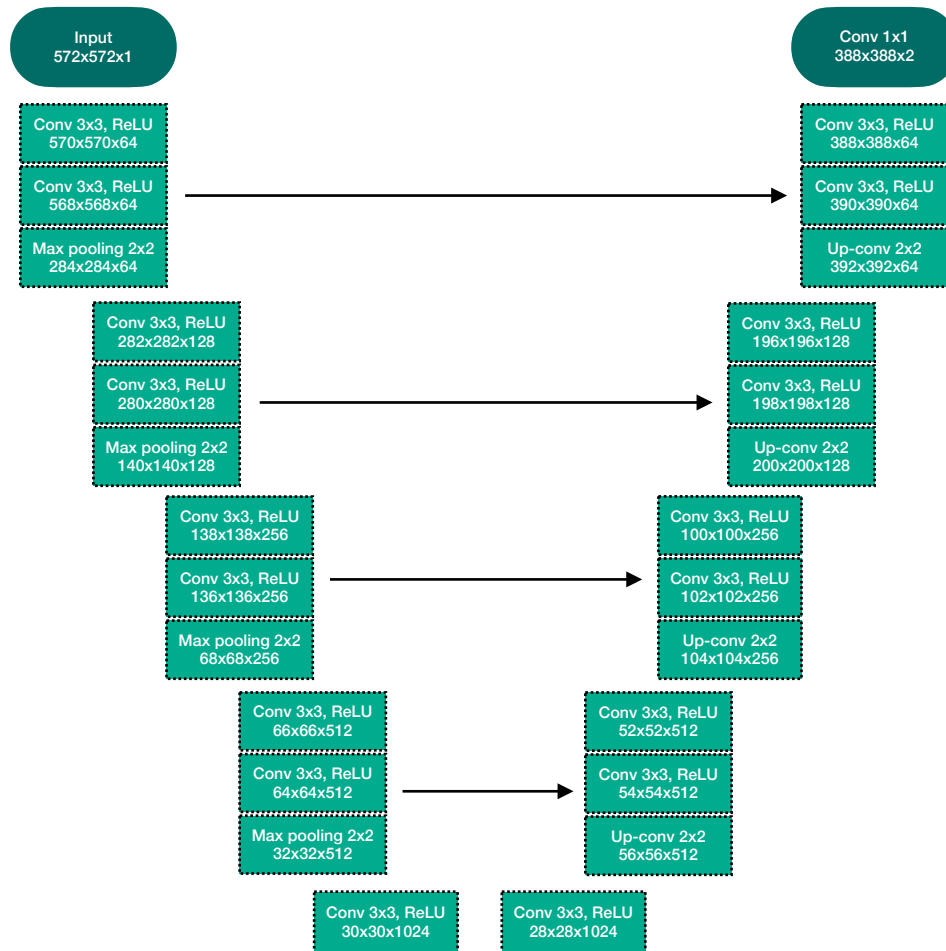


Figura 9

Arquitectura U-Net original planteada para segmentación de imágenes biomédicas utilizando una estrategia de tiles con overlap.

SegAN.

La arquitectura SegAN fue presentada como una red para la segmentación de imágenes médicas con un enfoque que usa aprendizaje adversarial profundo (Huang et al., 2017). Se compone de dos redes, una segmentadora y una crítica, que se entrenan de manera adversaria al igual que la arquitectura para generación de imágenes planteadas en el artículo original (Goodfellow et al., 2014). La primera red es la que genera la máscara de segmentación y la segunda es la que estima la probabilidad de que sea real o falsa, además de incorporar un ejemplo real de la máscara de segmentación. Ambas redes se enfrentan con objetivos distintos: por parte de la segmentadora es engañar a la red crítica y por parte de la red crítica es clasificar correctamente cuándo una imagen es real o falsa.

La red segmentadora que se muestra a continuación, tiene una estructura *encoder-decoder*, con función de activación Leaky ReLU en el *encoder* y ReLU en el *decoder*. Por otro lado, la red crítica que recibe las imágenes de la red segmentadora y las reales, se compone de tres bloques convolucionales con función de activación Leaky ReLU, para finalmente concatenar todos los saltos de conexión que se tienen entre los bloques. En la Figura 10, se muestra la estructura inicialmente propuesta para la segmentación semántica en imágenes médicas.

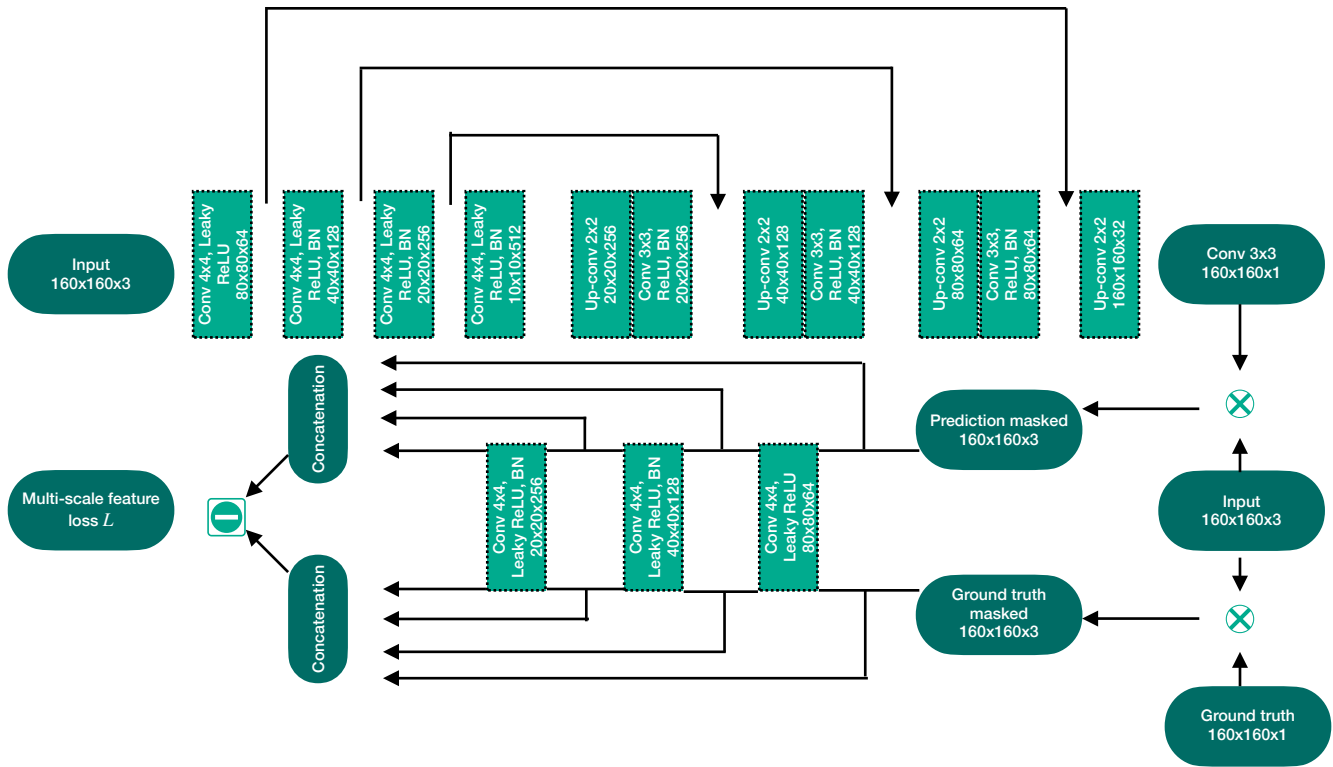


Figura 10

Arquitectura SegAN original planteada para segmentación de imágenes médicas.

TransUNet.

La TransUNet fue presentada como una arquitectura para segmentación semántica en imágenes médicas y que trajo consigo varios avances en la incorporación de transformadores a tareas más complejas como es la segmentación (Adeli et al., 2021). La estructura de la red está compuesta por dos bloques similares a la U-Net. El primero, un *encoder* que tiene un codificador de características de la arquitectura ResNet y un bloque de capas de transformadores. El segundo, un *decoder* que se compone de manera similar al utilizado por la U-Net, pero con componentes más complejos, llamados submultiplicadores en cascada. Además, de igual manera que la arquitectura antes mencionada, tiene conexiones de salto para la conservación de características intermedias que se van obteniendo. La idea que hay detrás de

la TransUnet es que al utilizar solamente transformadores en la arquitectura, a esta le cuesta realizar tareas más complejas como la segmentación. Sin embargo, al añadir la estructura de red neuronal convolucional de la U-Net se pueden aprender características más finas para realizar la segmentación. En la Figura 11, se muestra la estructura inicialmente propuesta para la segmentación de imágenes médicas.

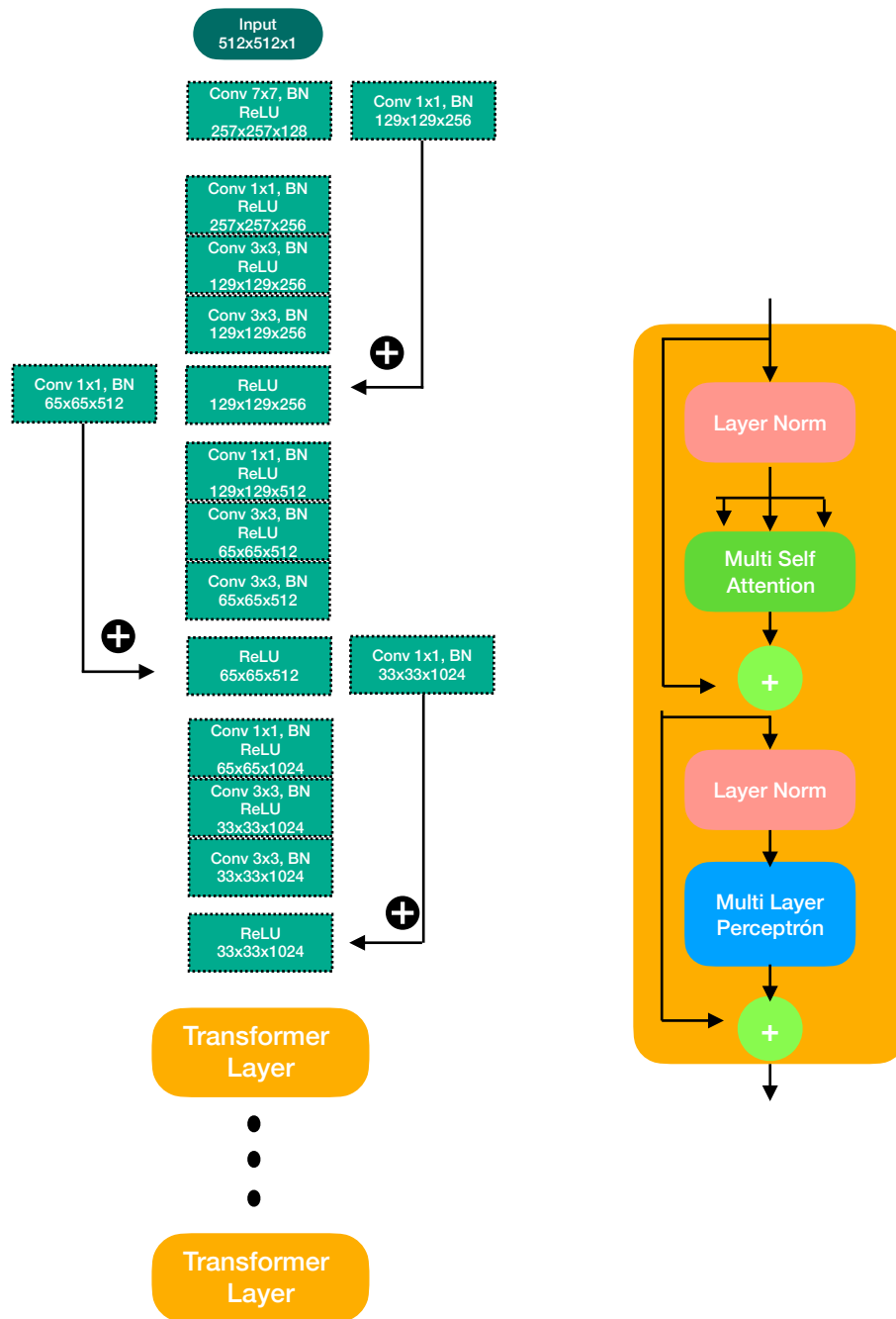


Figura 11
Encoder de arquitectura TransUNet.

Estado del arte

En la tarea de segmentación automática de la perfusión miocárdica en MRI cardíaca no existe un único set de datos estándar. Por ello, medir el avance del estado del arte adecuadamente se limita a las comparaciones realizadas por cada autor de los diferentes modelos existentes. Debido a esto, se realiza una revisión de los artículos más relevantes para establecer líneas de acción en la tarea antes mencionada.

A continuación, se presentan los diferentes enfoques utilizados para resolver el problema, partiendo desde redes neuronales convolucionales hasta aprendizaje adversarial profundo.

Redes neuronales convolucionales

Estas redes desde que se introdujeron al aprendizaje profundo, comenzaron a ser adaptadas y utilizadas en diferentes tareas asociadas a datos multimedia. Por ello, es importante analizar la literatura de las soluciones que tienen como idea base el uso de redes neuronales convolucionales.

La principal arquitectura utilizada es la U-Net, debido principalmente a sus resultados de vanguardia y adaptabilidad a diferentes problemas, una muestra de su uso en la segmentación de imágenes de perfusión cardíaca es en los artículos de [García et al. \(2021\)](#), [Brown et al. \(2020\)](#), [Arai et al. \(2020\)](#), [Choe et al. \(2019\)](#) y [Breeuwer et al. \(2019\)](#). Además, es importante notar que diferentes artículos que han utilizado la U-Net proponen mejoras importantes que se adaptan adecuadamente al problema de segmentación en perfusión cardíaca. Por ejemplo, el uso de aprendizaje por transferencia introducido en [García et al. \(2021\)](#) que demuestra que el pre-entrenamiento en imágenes de cine cardíaco contribuye a lograr que el modelo converja a mejores resultados. Así también, el artículo de [Brown et al. \(2020\)](#) que

introduce la corrección de movimiento en los fotogramas como procesamiento previo a ser segmentada por la U-Net.

Por otro lado, ideas más complejas como la introducción del factor temporal en la segmentación que es un enfoque interesante debido a la naturaleza secuencial del problema planteada en [Arai et al. \(2020\)](#). Por último, se tienen las novedades introducidas por [Choe et al. \(2019\)](#) y [Breeuwer et al. \(2019\)](#), que utilizan novedosos métodos como el “dropout” de Monte Carlo y la detección de una caja que delimita la zona a segmentar, respectivamente.

En general, los artículos que utilizan redes neuronales convolucionales como base y en específico la U-Net son la gran mayoría del estado del arte en segmentación de la perfusión cardíaca en resonancia magnética cardíaca. Sin embargo, es necesario introducir técnicas para que dicha segmentación logre mejores resultados.

Redes neuronales recurrentes

Previo a la aparición de los transformadores, las redes neuronales recurrentes eran el estado del arte en tareas que tenían una componente secuencial. Por lo que estudiar los artículos que utilizan dichas redes es importante para comprender estos enfoques secuenciales.

El único artículo en el estado del arte donde se propone directamente el uso de redes neuronales recurrentes es [Fei et al. \(2021\)](#). Este plantea una arquitectura similar a la U-Net pero que introduce un *decoder* con redes neuronales recurrentes para comprender el inherente cambio temporal que existe en las imágenes de perfusión cardíaca. En particular, se utiliza el tipo de arquitectura recurrente que es la de *Long Short-Term Memory* o LSTM.

Aprendizaje adversarial profundo

El aprendizaje adversarial profundo es una línea de arquitecturas que ha sido aplicada a diferentes dominios, y en particular, es utilizada en la segmentación del miocardio. En el caso específico del artículo de [Wang et al. \(2021\)](#), se utiliza un tipo de imágenes distinto a las

disponibles para la presente investigación. No obstante, se considera de igual manera relevante estudiar el estado del arte relacionado a la aplicación de dichos métodos, por su novedosa introducción en la tarea de segmentación de imágenes médicas.

En dicho artículo introducen una estructura segmentador-crítico que en su función de pérdida introduce un factor α que le quita importancia a la función de pérdida adversarial para el segmentador. De esta manera, se puede lograr un entrenamiento más estable ofreciendo resultados de vanguardia.

Otros enfoques

En la literatura también se han explorado otros métodos menos convencionales que se alejan levemente lo que es el aprendizaje profundo. Uno de ellos es el presentado por Jacobs et al. (2021) que corresponde a un método automático de segmentación mediante algoritmos matemáticos de detección de bordes y algoritmos de procesamiento de las imágenes que reconstruyen la zona detectada. Sin embargo, este método se aleja de los enfoques más prácticos utilizados para resolver el problema, por lo que no se contempla como información relevante para plantear la metodología.

Metodología

En base a lo analizado en el Marco teórico y el Estado del arte, se propone la siguiente metodología que involucra desde el procesamiento previo de las imágenes hasta la experimentación con diferentes modelos de aprendizaje profundo. De esta manera, se espera comparar y obtener los mejores resultados para la segmentación automática de la perfusión miocárdica en imágenes de resonancia magnética cardíaca.

Revisión del etiquetado de las imágenes

Una de las mayores dificultades en el proceso de investigación está relacionado al etiquetado de las imágenes. Esto debido, principalmente, a la diversidad de criterios de segmentación que utilizan los profesionales encargados de la tarea de delimitar los objetos interés en las imágenes de resonancia magnética. Por ello, los algoritmos de segmentación pueden tener resultados limitados a causa de los diferentes criterios utilizados. Por lo tanto, se decide estandarizar el etiquetado en base a la opinión de expertos, criterios utilizados en otros set de datos de imágenes de perfusión y lineamientos establecidos por artículos de investigación relacionados al tema. En particular, considerando el análisis de la segmentación realizada en Kellman y Wollny (2014), criterios definidos en Bremerich et al. (2013) y recomendaciones realizadas previamente por expertos del Hospital Clínico de la Universidad de Chile. Se proponen los siguientes pasos para el correcto estudio y etiquetado de las imágenes de perfusión cardíaca.

Criterios planteados:

1. Se plantea utilizar el *software* de segmentación de imágenes médicas ITK-SNAP, debido a la disponibilidad gratuita y capacidad de etiquetado. Se configura la segmentación en base a dibujar polígonos en los bordes de los objetos de interés.

2. Aplicar los mismos ajustes de contraste, brillo y ventana en todas las imágenes de las series. En el caso de P-UCH, se utilizan las configuraciones automáticas del software ITK-SNAP.

3. Se propone segmentar las líneas correspondientes al epicardio y endocardio, que son las que rodean el miocardio. De esta manera, se puede realizar el análisis sobre las tres clases de interés en la detección de enfermedades cardiovasculares.

4. En base a lo observado en Kellman y Wollny (2014) se propone segmentar la línea del endocardio por el cambio de color más externo que existe en algunas imágenes más complejas, lo que implica incluir algunas zonas grises que están directamente conectadas a las zonas más blancas externas a las tres clases de interés.

5. En base a lo observado en Kellman y Wollny (2014) se propone segmentar la línea del epicardio por el cambio de color de tonos grises a claros en la zona que delimita con el ventrículo izquierdo.

6. En base a lo observado en Kellman y Wollny (2014) se propone segmentar la línea del epicardio por el cambio de color de tono oscuro a claros que existe en el lado opuesto al ventrículo izquierdo.

7. En base a lo observado en Kellman y Wollny (2014) se propone segmentar el epicardio realizando una progresión aproximada del ancho y curvatura del miocardio en las zonas donde no existe un cambio de tonalidad visible de la línea del epicardio y partes externas.

Debido a limitaciones de tiempo y recursos, se revisa y corrige solamente la clase relacionada al endocardio, el cual fue un proceso que tardó 2 meses en ser revisado. Pero, de igual forma se proponen indicaciones para el borde del epicardio a pesar de que se mantienen las etiquetas realizadas previamente.

Procesamiento previo

Para el procesamiento previo se utiliza un algoritmo de detección de la región de interés basado en algoritmos de saliencia. En particular, se propone utilizar el algoritmo *Graph-Based Visual Saliency* (GBVS), que a pesar de no ser un algoritmo relativamente nuevo, logra resultados muy satisfactorios en la tarea propuesta y permite reducir el tamaño de la imagen al combinarlo con la idea propuesta en [Akbari et al. \(2018\)](#) para generar un cuadro delimitador sobre los objetos salientes detectados.

Graph-Based Visual Saliency (GBVS)

En el trabajo previo de [García et al. \(2021\)](#) con las imágenes de perfusión cardiaca se plantea utilizar un algoritmo de saliencia para obtener imágenes más pequeñas y centradas en el objeto de interés. De esta manera, con las imágenes de 128 x 128 píxeles se espera tener un entrenamiento mucho más rápido comparado a utilizar las imágenes completas de 256 x 256 píxeles, al contar con matrices más pequeñas que involucra menos procesamiento para las redes neuronales artificiales. En base a lo planteado por [Akbari et al. \(2018\)](#), se propone el siguiente método de cinco pasos para la extracción de la región de interés en las imágenes de perfusión.

1. Diferencia absoluta de cada dos fotogramas consecutivos: el resultado especifica la variación de intensidades entre dos fotogramas consecutivos, que incluye el movimiento del corazón junto con mucho ruido.
2. Suma de todas las diferencias absolutas: debido a la gran cantidad de ruido en cada imagen de diferencia absoluta, el corazón no es una parte muy explicativa en la diferencia absoluta resultante. Por lo tanto, la suma de todas las diferencias absolutas relacionadas con una secuencia hace que el corazón sea más destacado en la suma resultante.
3. Detección de saliencia: se utiliza el algoritmo GBVS planteado para calcular el mapa de saliencia sobre la suma de las diferencias absolutas.

4. Umbralización: se realiza un umbral en el mapa de saliencia y se extrae el 90% de las partes más salientes.
5. Aplicación del cuadro delimitador: se aplica un cuadrado delimitador de tamaño fijo alrededor de la parte saliente, con el que se obtiene una imagen de 128 x 128 píxeles.

Es importante destacar que el proceso para la detección de la región de interés solo se aplica sobre las imágenes de perfusión, ya que, son el objeto de la presente investigación, no así en tanto las imágenes de cine que se utilizan solamente para pre-entrenamiento.

Modelos

Los modelos que se plantean para resolver la tarea de segmentación semántica en las imágenes de perfusión cardíaca son modificaciones de las arquitecturas originales planteadas en los artículos de investigación de referencia. En ellos, se modifican aspectos como la entrada, salida, cantidad de bloques, funciones de activación, etc. Además, para el correcto control del funcionamiento de las arquitecturas, se plantean los diagramas con todas las partes del modelo y el tamaño de las salidas después de cada componente. De esta manera, se puede construir y modificar rápidamente una arquitectura en virtud de los resultados que se van obteniendo en los experimentos y lo que se observa a continuación, es el resultado del proceso. Por último, es importante notar que las arquitecturas propuestas a excepción de la U-Net no han sido aplicadas al caso específico de perfusión cardíaca. Por lo que, es de interés analizar sus resultados en comparación al modelo base que es la U-Net y estudiar la adición de nuevas líneas de investigación, si los resultados no logran ser concluyentes.

U-Net modificada

En la Figura 12, se presenta una arquitectura modificada de la U-Net original, que se adapta mejor a la complejidad del problema. De esta manera, se presenta el diagrama de la arquitectura propuesta con los diferentes componentes que la diseñan, como la capa de entrada y de salida, los bloques basados en redes neuronales convolucionales, los saltos de conexión para la conservación de las características intermedias, entre otros componentes.

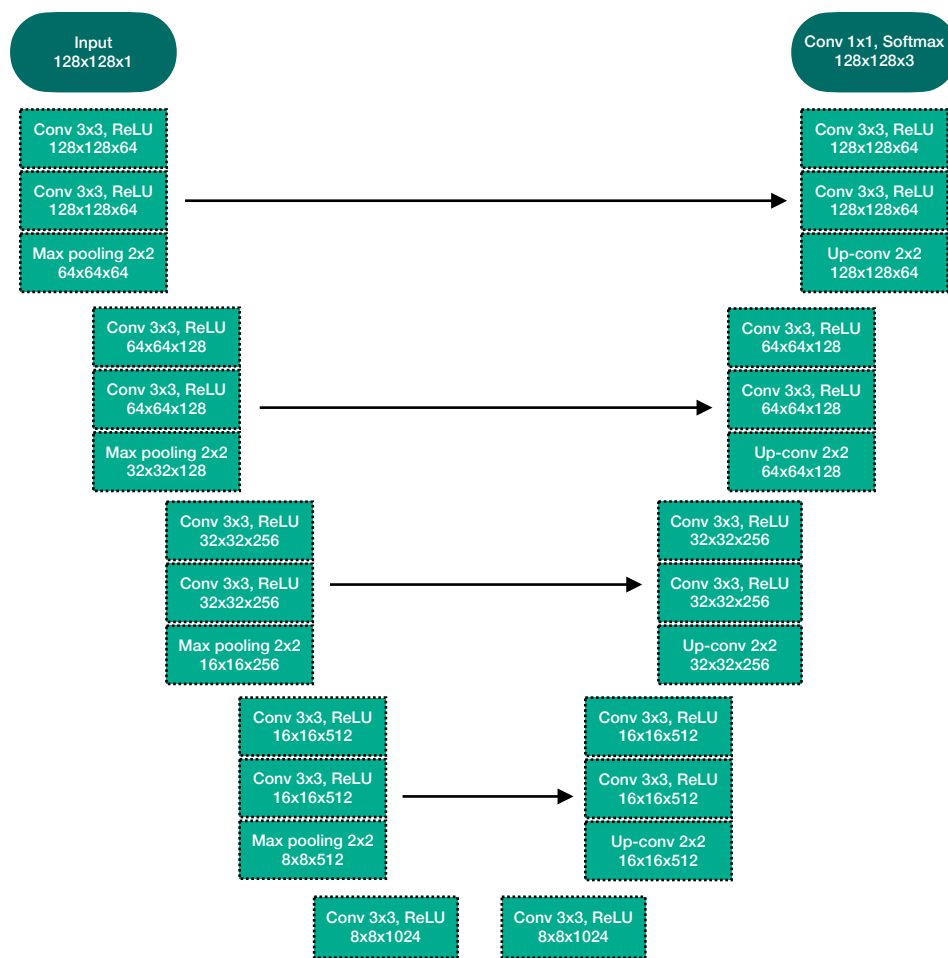


Figura 12
Arquitectura U-Net modificada.

A continuación, se muestra la función de pérdida adversarial planteada en el artículo (Wang et al., 2021), que consiste en una función mínimo-máximo que se utiliza en la red discriminadora y en la red segmentadora. Además, se muestra la función de pérdida utilizada en la red segmentadora.

$$(6) \quad \min_{Seg} \max_{Disc} Adv_{loss}(Seg, Disc) = \frac{1}{N} \sum_{n=1}^N l_{mae}(Disc(x_n * Seg(x_n)), Disc(x_n * y_n))$$

Donde N es el número de imágenes de entrenamiento, x_n la imagen y la etiqueta real y_n , l_{mae} es el error absoluto promedio o la distancia L_1 , $x_n * Seg(x_n)$ es la máscara de la imagen con las predicciones de la red segmentadora y $x_n * y_n$ la máscara de la imagen con la etiqueta real. Además, l_{mae} se define como la Ecuación 7.

$$(7) \quad l_{mae}(Disc(x), Disc(x')) = \frac{1}{L} \sum_{i=1}^L ||Disc_i(x) - Disc_i(x')||$$

Con L el número de capas en la red discriminadora y $Disc_i(x)$ es el mapa de características obtenido de la imagen x en la capa i -ésima de la misma red.

$$(8) \quad loss = dice_{loss} + \alpha * Adv_{loss}$$

TransUNet

Se plantea una arquitectura modificada de la TransUNet explicada anteriormente, donde se varía principalmente en características leves o hiperparámetros que modifican el tamaño de la red. Se propone utilizar una red neuronal basada en transformadores como la TransUNet

porque de esta manera se puede analizar la imagen como una secuencia de parches y comprender cómo se relacionan las diferentes componentes dentro de una misma imagen. Además, es una arquitectura que ha demostrado excelentes resultados en imágenes de cine cardíaca, por lo que es intuitivo esperar resultados de buen nivel en imágenes de perfusión cardíaca. Finalmente, se muestra el diagrama planteado para la TransUNet modificada, donde se muestran sus principales componentes y tamaños de salida de cada bloque.

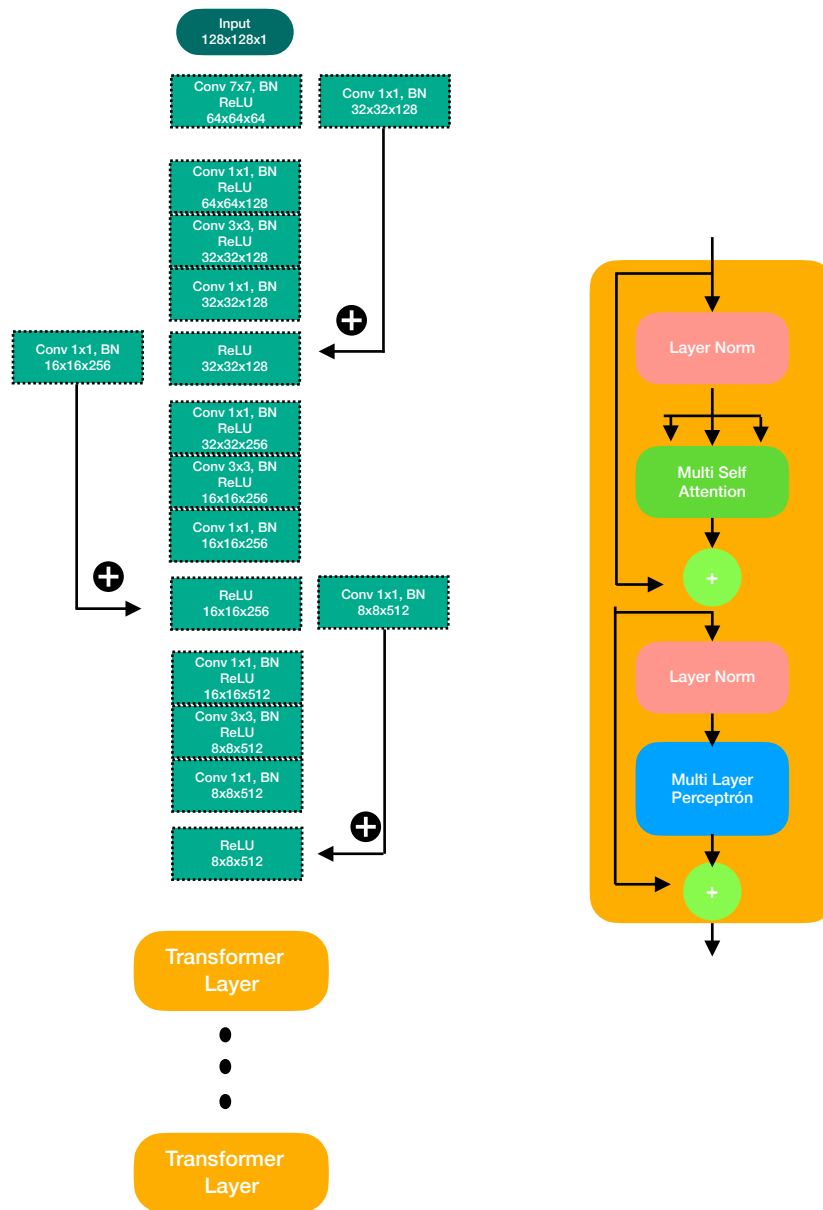


Figura 14
Encoder de la arquitectura TransUNet.

Métricas

Una parte importante de la utilización de modelos de aprendizaje automático son las métricas de evaluación, de esta manera, se puede estimar cuán bueno es el modelo para realizar la tarea que tiene como objetivo. En el caso de la segmentación semántica, algunas de las métricas más utilizadas son el coeficiente de similitud dice (DSC) y la intersección sobre unión (IoU).

Coefficiente de Similitud Dice

El coeficiente de similitud Dice o índice de Sorensen-Dice es del tipo basada en F y una de las más utilizadas para evaluar el rendimiento en visión computacional. El tipo de métricas F se calculan en base a la sensibilidad y precisión de una predicción, con lo que se calcula un puntaje asociado al solapado entre la máscara predicha y la de verdad. En particular, la métrica Dice, se define como la media armónica entre la sensibilidad y la precisión. A continuación, se muestra la Ecuación 9 de la métrica Dice, con TP: verdaderos positivos, FP: falsos positivos y FN: falsos negativos.

$$(9) \quad DSC = \frac{2TP}{2TP + FP + FN}$$

Intersección sobre Unión

La métrica intersección sobre unión o IoU sirve para medir cuan solapadas están dos máscaras de segmentación y es similar al coeficiente de similitud Dice, ambos son del tipo pasado en F. Sin embargo, la diferencia entre ambas métricas es que IoU penaliza más la infra-segmentación y la sobre-segmentación, con TP: verdaderos positivos, FP: falsos positivos y FN: falsos negativos.

(10)
$$IoU = \frac{TP}{TP + FP + FN}$$

Otras técnicas

A continuación, se describen otras técnicas utilizadas con distintos propósitos, como lo puede ser mejorar el rendimiento de los modelos anteriores o para contribuir a la estabilidad del entrenamiento. Las técnicas descritas son ampliamente utilizadas en la literatura y es una práctica usual en el aprendizaje automático, debido a su alta efectividad para apoyar a algoritmos complejos a obtener mejores resultados o tener mayor estabilidad. De esta manera, se decide utilizar las técnicas de aumento de datos, aprendizaje por transferencia y otras transformaciones de las imágenes.

Transformaciones previas

Para estandarizar las imágenes previas a ser entregadas a los modelos de aprendizaje profundo, se realizan ciertas modificaciones fijas que no cuentan como transformaciones momentáneas como lo son las realizadas para aumentar datos. A continuación, se entregan más detalles al respecto.

- Escala de valores: es necesario transformar la escala de los valores en cada píxel de las imágenes, debido a que desde el software de imágenes médicas vienen con una escala completamente diferente, con valores que superan la escala estándar utilizada en visión computacional.
- Recorte de valores extremos: los valores de los píxeles se recortan al percentil 99th para eliminar los valores extremos.
- Relleno de ceros: para que todas las imágenes sean del mismo tamaño, cuando es necesario se realiza un relleno en el exterior de ceros para obtener la imagen estándar de 256 x 256 píxeles.

Aumento de datos

Se utilizan diferentes transformaciones de las imágenes para aumentar artificialmente la cantidad de ejemplos que los algoritmos utilizan para entrenarse y de esta manera, sean más robustos ante ejemplos nuevos y variados de los datos de entrada. A continuación, se describen brevemente algunas de las transformaciones utilizadas.

- Rotación: se rotan las imágenes y la máscara con las anotaciones de manera aleatoria en un ángulo entre -360° y 360° , con probabilidad de 50% de aplicarse o no.

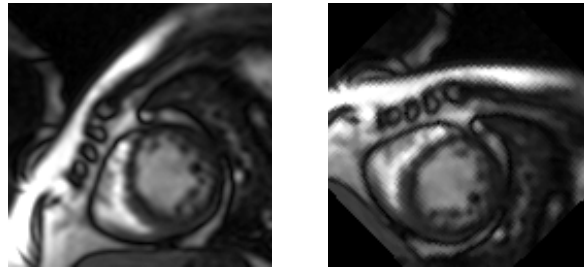


Figura 15

A la izquierda la imagen original y a la derecha la imagen con rotación en un ángulo aleatorio.

- Voltar verticalmente: se voltean las imágenes y la máscara con las anotaciones de forma vertical, con probabilidad de 50% de aplicarse o no.

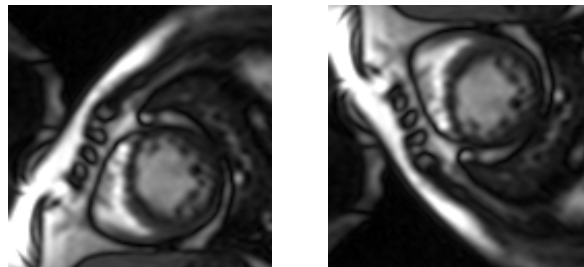


Figura 16

A la izquierda la imagen original y a la derecha la imagen volteada verticalmente.

- Voltear horizontalmente: se voltean las imágenes y la máscara con las anotaciones de forma horizontal, con probabilidad de 50% de aplicarse o no.

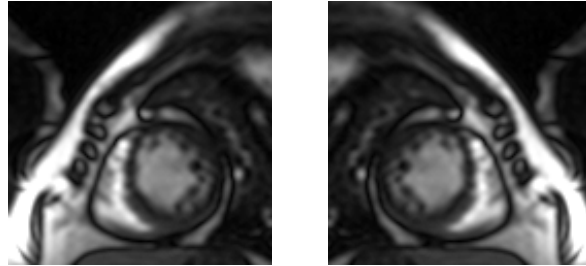


Figura 17

A la izquierda la imagen original y a la derecha la imagen volteada horizontalmente.

- Ajuste de brillo: se ajusta el brillo de las imágenes (no se considera la máscara) con un factor de brillo obtenido de forma aleatoria uniforme entre 0.5 y 2, con probabilidad de 50% de aplicarse o no

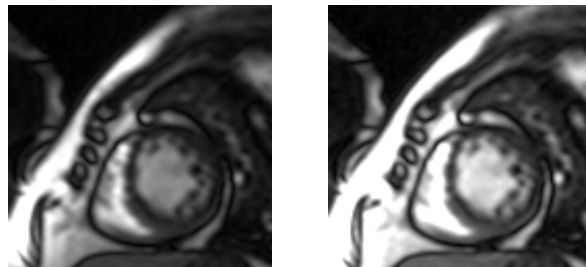


Figura 18

A la izquierda la imagen original y a la derecha la imagen con cambio en el brillo aleatorio.

- Ajuste de contraste: se ajusta el contraste de las imágenes (no se considera la máscara) con un factor de contraste obtenido de forma aleatoria uniforme entre 0.5 y 2, con probabilidad de 50% de aplicarse o no.

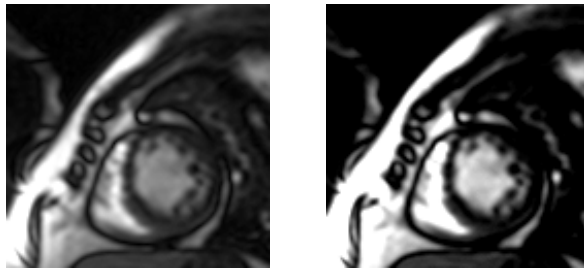


Figura 19

A la izquierda la imagen original y a la derecha la imagen con cambio en el contraste aleatorio.

Aprendizaje por transferencia

Una de las técnicas que más han sido utilizadas en la investigación es el aprendizaje por transferencia, desde aplicaciones en imágenes donde se entrenan los modelos previamente en enormes conjuntos de datos y luego se utiliza ese aprendizaje a priori para estabilizar o mejorar el rendimiento en tareas más específicas. De esta manera, se propone hacer aprendizaje por transferencia desde el pre-entrenamiento en imágenes de cine y luego, un ajuste fino en el conjunto de imágenes de interés, que son las imágenes de perfusión. El objetivo en este caso específico es tener un algoritmo que “comprende” de mejor manera la estructura de las imágenes previamente, ya que, las imágenes de cine son en esencia bastante similares a las imágenes de perfusión, por lo que realizar un entrenamiento previo en ellas contribuye a que aprenda de manera más estable y se pueda llegar de mejor manera a un óptimo local o global.

Regularización

Las técnicas de regularización son una parte muy importante para el adecuado entrenamiento y evitar el sobre ajuste en los modelos, por ello, se propone utilizar algunas técnicas para poder evitar dichos problemas y que los resultados sean los adecuados. A

continuación, se describen algunas de las técnicas de regularización utilizadas en la implementación de los modelos propuestos.

- Decaimiento de los pesos: con esta técnica se penalizan los pesos grandes en función de sus valores al cuadrado, lo cual implica añadir una penalización en la función de coste. La penalización mantiene los pesos pequeños, evitando que la red utilice pesos que no necesita.
- Normalización del lote: consiste en añadir una capa de normalización en los bloques de la red neuronal, en el que se normalizan las activaciones de salida por la media y varianza de los lotes de entrada. Además, se añaden dos parámetros que son un sesgo que se suma y una constante que se multiplica por cada valor de activación, lo cual permite que el rango de entrada escale de mejor manera al rango de salida, y reduce las oscilaciones de la función de costos. Finalmente, debido a esto se puede aumentar la tasa de aprendizaje sin tanto riesgo de quedar atrapado en un mínimo local y la convergencia hasta un mínimo global se produce mucho más rápido, pudiendo entrenar los modelos por menos épocas.
- *Dropout*: consiste en desactivar aleatoriamente un porcentaje de las neuronas definido previamente. De esta manera, se evita que las neuronas memoricen partes de la entrada y por ello, se tenga un sobre ajuste.

Resultados

Bases de datos

Para el entrenamiento de los modelos y posterior a la revisión del set de datos de perfusión cardíaca del Hospital Clínico de la Universidad de Chile, se cuenta con tres bases de datos para los objetivos de pre-entrenamiento, entrenamiento en los datos de interés y validación externa en datos del mismo tipo. En la Figura 20, se muestran las bases de datos disponibles, que son ACDC correspondientes a imágenes de Cine cardíaca, P-UCH que son de perfusión cardíaca y P-GIGA que es la base de datos externa de perfusión cardíaca del artículo de [Kellman y Wollny \(2014\)](#).

	ACDC	P-UCH	P-GIGA
Tipo	Cine	Perfusión	Perfusión
Sujetos	100	88	9
Fotogramas totales	1808	6744	505
Fotogramas de entrenamiento	1538	5131	No aplica
Fotogramas de validación	270	506	No aplica
Fotogramas de prueba	No aplica	1107	505

Figura 20
Comparativa bases de datos disponibles.

Detalles del entrenamiento

A continuación, se muestran los hiperparámetros estándar utilizados para los diferentes modelos, como lo son el tamaño del lote, la tasa de aprendizaje, el optimizador utilizado y cantidad de iteraciones.

	U-Net	U-Net + Aprendizaje Adversarial Profundo	TransUNet
Optimizador	Adam	Adam	Adam
Cantidad de épocas	500**	500**	500**
Tamaño del lote	36	24	36
Función de pérdida	Dice loss	Dice loss + α * Adversarial loss	Dice loss

Figura 21

Comparativa detalles de entrenamiento arquitecturas. ** La cantidad de iteraciones fue variada porque los experimentos se detuvieron de forma temprana al notar sobreentrenamiento.

En el entrenamiento, se realizó la validación del modelo cada época donde se computa la función de pérdida de validación y se guarda el estado actual del modelo en caso de que el valor de la pérdida sea inferior al resultado del mejor modelo encontrado anteriormente, para luego actualizar el mejor modelo con el encontrado recientemente.

U-Net

Para obtener los mejores resultados con la arquitectura U-Net, se realiza una exhaustiva experimentación con diferentes configuraciones de hiperparámetros de la arquitectura, analizando el comportamiento en los resultados de las métricas, convergencia y complejidad del entrenamiento. Finalmente, después de los resultados obtenidos en la Figura 22 en el anexo se llega a la siguiente configuración para el entrenamiento de los modelos principales.

Hiperparámetros de la arquitectura	Features iniciales	Tasa de aprendizaje	Tamaño del lote	Valor del decaimiento del peso
Valores	64	0.0001	36	0.001

Figura 22

Búsqueda de hiperparámetros específicos de la arquitectura.

Para obtener resultados fiables de la capacidad de la arquitectura para predecir sobre la tarea especificada, se realizan 10 entrenamientos del modelo sin pre-entrenamiento en imágenes de cine y se promedian las métricas obtenidas. Además, se realizan 10 entrenamientos con pre-entrenamiento en imágenes de cine. A continuación, se presenta la Figura 23 con las métricas promedio obtenidas sobre los datos de prueba al realizar los experimentos con pre-entrenamiento y sin pre-entrenamiento.

Pre-entrenamiento	Dice miocardio (%) ± std (%)	Dice endocardio (%) ± std (%)	Dice epicardio (%) ± std (%)	IoU miocardio (%) ± std (%)	IoU endocardio (%) ± std (%)	IoU epicardio (%) ± std (%)
✘	82,51 ± 0,51	92,3 ± 0,090	93,5 ± 0,17	83,98 ± 0,40	92,36 ± 0,086	93,28 ± 0,17
✔	82,48 ± 0,29	92,18 ± 0,24	93,53 ± 0,096	83,94 ± 0,23	92,24 ± 0,22	93,3 ± 0,083

Figura 23

Resultados de modelos principales sobre el set de datos P-UCH.

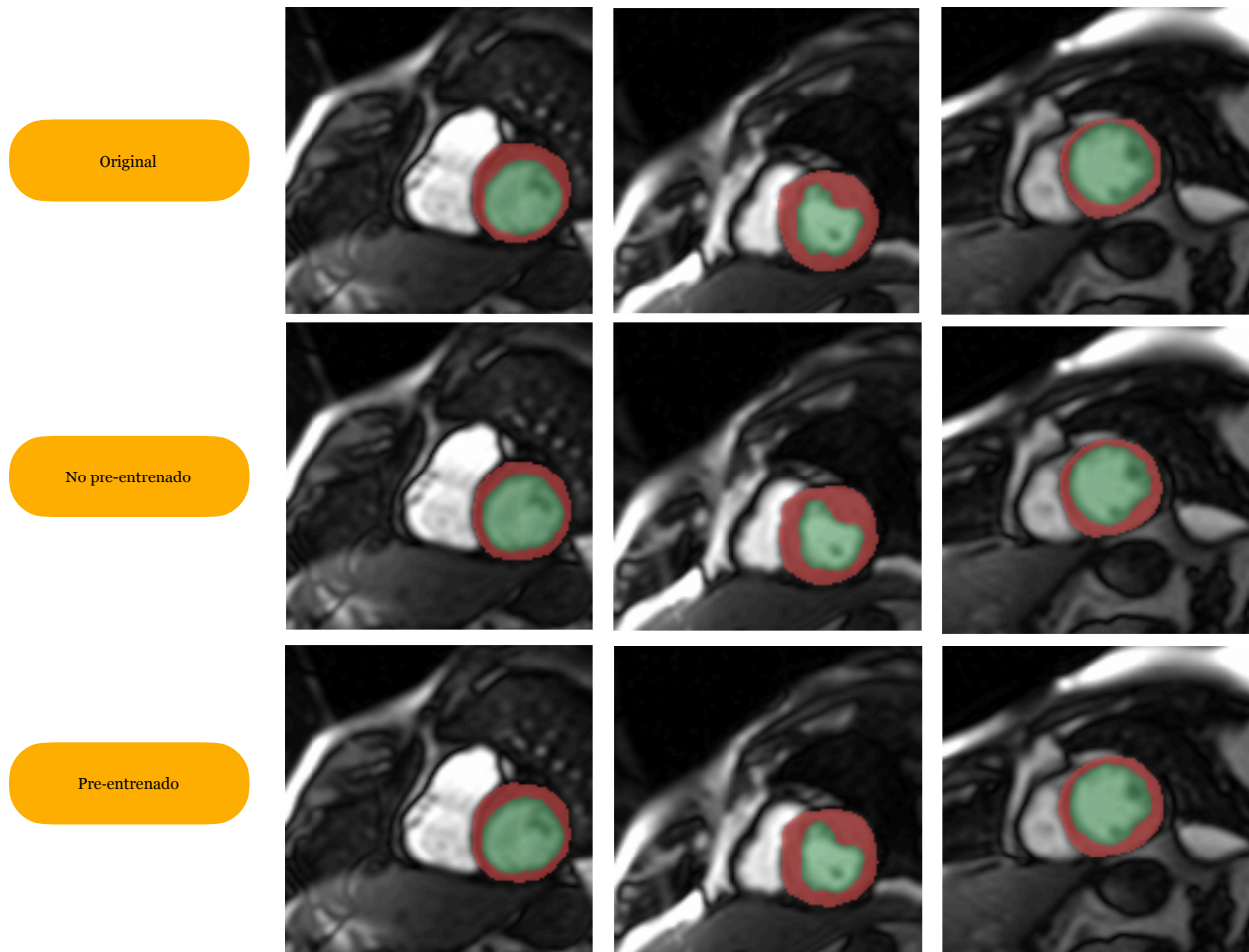


Figura 24

Comparativa resultados de segmentación u-net, con tres imágenes de ejemplo del set de datos de testeo. Primera fila, segmentación manual, segunda fila es una U-Net no pre-entrenada y tercera fila es la U-Net pre-entrenada en imágenes de cine.

A continuación, se disponen tableros donde se pueden observar resultados de las segmentaciones, las métricas por cada modelo, los entrenamientos y la validación de los experimentos con pre-entrenamiento y los sin pre-entrenamiento.

- Sin pre-entrenamiento: [reporte 1](#)
- Con pre-entrenamiento: [reporte 2](#)

U-Net + Aprendizaje Adversarial Profundo

Para obtener los mejores resultados con la arquitectura SegAN, se va a utilizar la arquitectura U-Net como base y la configuración planteada anteriormente, con la excepción del tamaño del lote que se reduce a 24 por limitaciones en la memoria RAM y el tipo de entrenamiento. Por otro lado, es necesario definir adecuadamente un α que multiplica a la función de pérdida adversarial en la ecuación propuesta en la ecuación descrita en la metodología. De esta manera, se realiza una exhaustiva experimentación con diferentes valores de α , y en la Figura 25, Figura 26 y Figura 27 se puede ver el comportamiento de las métricas con respecto a ese hiperparámetro.

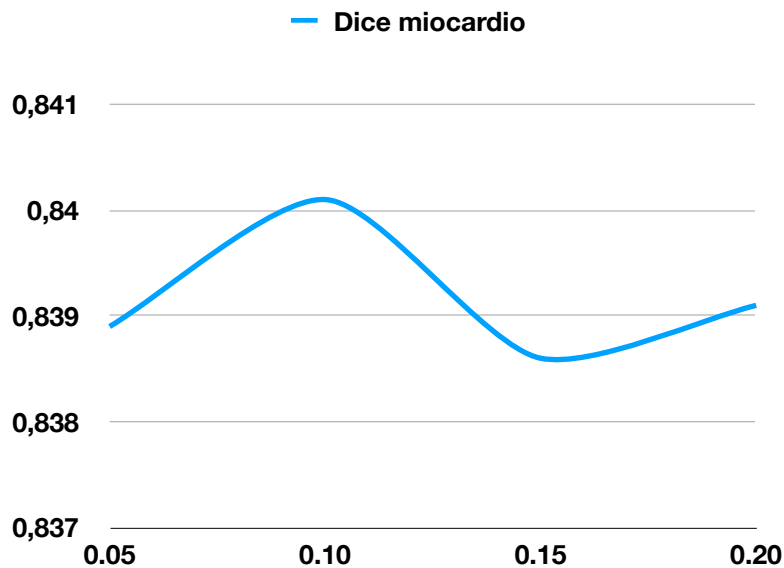


Figura 25
Comportamiento de la métrica del miocardio
con respecto al factor α .

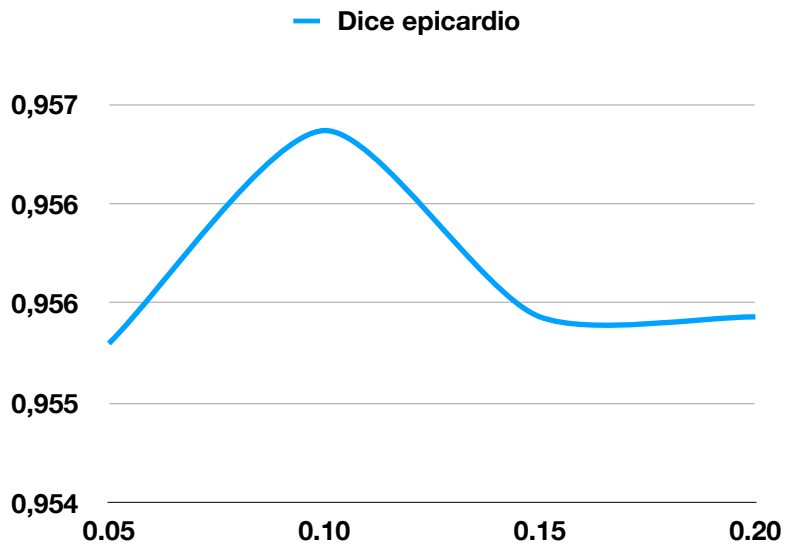


Figura 26
Comportamiento de la métrica del epicardio con respecto al factor α .

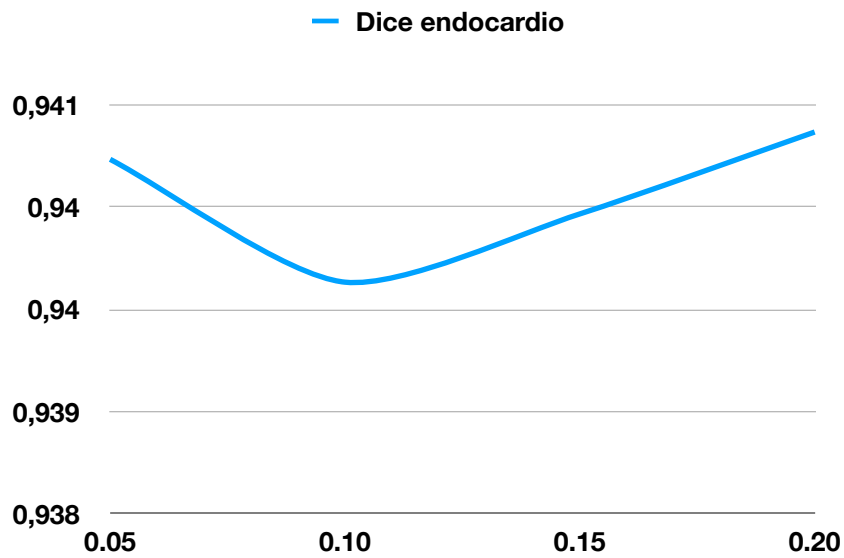


Figura 27
Comportamiento de la métrica del endocardio con respecto al factor α .

En base a los gráficos y que la principal clase de interés es el miocardio, se decide utilizar un α de 0.1, que si bien es el que ofrece mejores resultados en las pruebas realizadas, en realidad las diferencias entre los valores de α planteados es muy baja.

Para obtener resultados fiables de la capacidad de la arquitectura para predecir sobre la tarea especificada, se realizan 10 entrenamientos del modelo sin pre-entrenamiento en imágenes de cine y se promedian las métricas obtenidas. Además, se realizan 10 entrenamientos con pre-entrenamiento en imágenes de cine. A continuación, se presenta la Figura 28 con las métricas promedio obtenidas sobre los datos de prueba al realizar los experimentos con pre-entrenamiento y sin pre-entrenamiento.

Pre-entrenamiento	Dice miocardio (%) ± std (%)	Dice endocardio (%) ± std (%)	Dice epicardio (%) ± std (%)	IoU miocardio (%) ± std (%)	IoU endocardio (%) ± std (%)	IoU epicardio (%) ± std (%)
✘	83,98 ± 0,34	94,07 ± 0,11	95,55 ± 0,12	85,58 ± 0,27	94,16 ± 0,10	95,32 ± 0,12
✔	84,08 ± 0,22	94,1 ± 0,14	95,57 ± 0,11	85,65 ± 0,17	94,18 ± 0,13	95,34 ± 0,11

Figura 28

Resultados de modelos principales sobre el set de datos P-UCH.

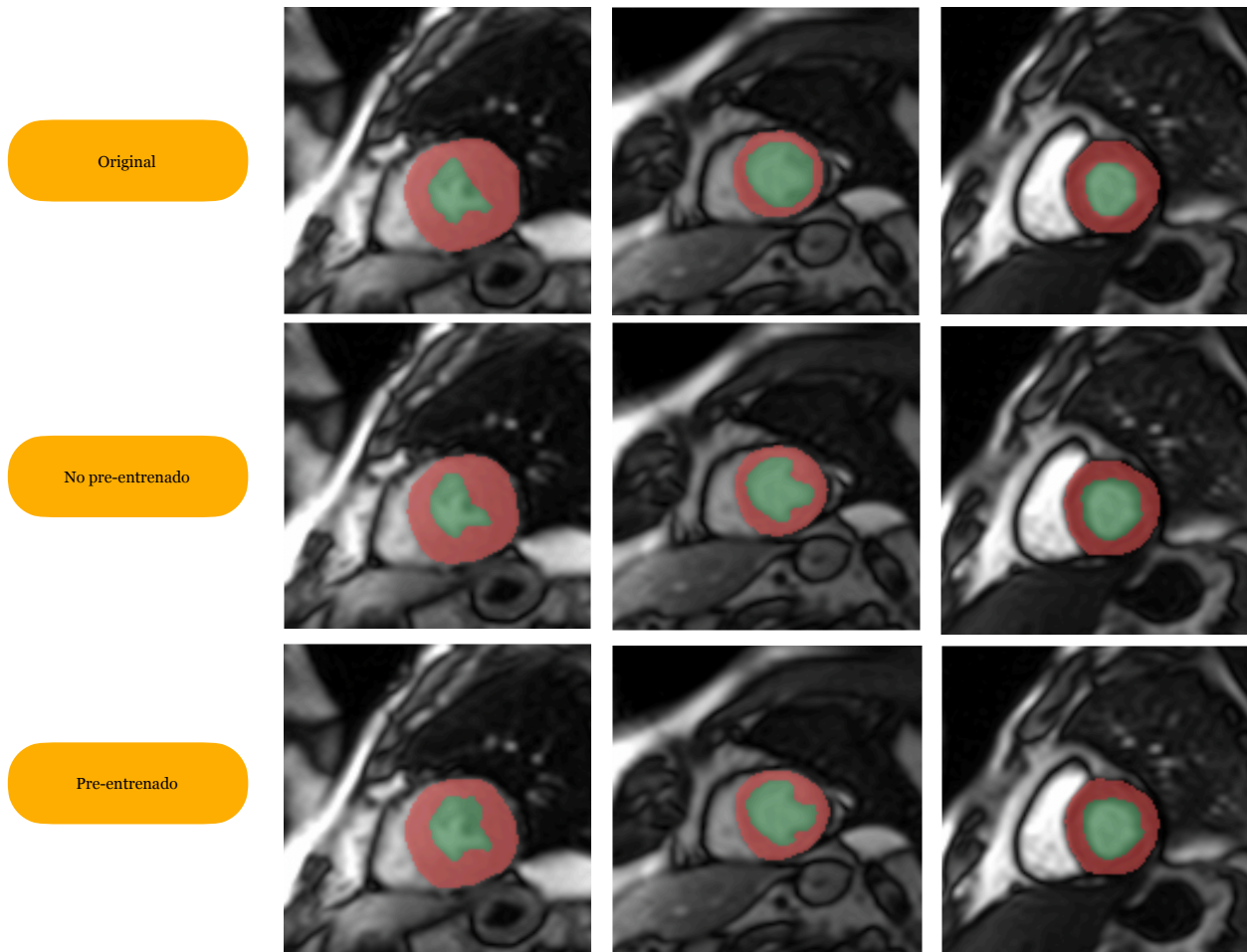


Figura 29

Comparativa resultados de segmentación u-net, con tres imágenes de ejemplo del set de datos de testeo. Primera fila, segmentación manual, segunda fila es una SegAN no pre-entrenada y tercera fila es la SegAN pre-entrenada en imágenes de cine.

A continuación, se disponen tableros donde se pueden observar resultados de las segmentaciones, las métricas por cada modelo, los entrenamientos y la validación de los experimentos con pre-entrenamiento y los sin pre-entrenamiento.

- Sin pre-entrenamiento: [reporte 1](#)
- Con pre-entrenamiento: [reporte 2](#)

TransUNet

Para obtener los mejores resultados con la TransUNet, en base a lo planteado en el artículo original se van a realizar experimentos con la mejor red que se obtuvo en dicho texto, que ya fue probada en imágenes de cine cardíaco y obtuvo resultados de vanguardia. De esta manera, se realizan 10 experimentos del modelo sin pre-entrenamiento en imágenes de cine y se promedian las métricas obtenidas. Además, se realizan 10 entrenamientos con pre-entrenamiento en imágenes de cine. A continuación, se presenta la Figura 30 con las métricas promedio obtenidas sobre los datos de prueba al realizar los experimentos con pre-entrenamiento y sin pre-entrenamiento.

Pre-entrenamiento	Dice miocardio (%) ± std (%)	Dice endocardio (%) ± std (%)	Dice epicardio (%) ± std (%)	IoU miocardio (%) ± std (%)	IoU endocardio (%) ± std (%)	IoU epicardio (%) ± std (%)
✗	83,21 ± 0,59	93,75 ± 0,20	95,3 ± 0,21	85,01 ± 0,44	93,88 ± 0,17	95,08 ± 0,20
✓	83,55 ± 0,34	93,89 ± 0,15	95,42 ± 0,11	85,27 ± 0,26	93,99 ± 0,14	95,19 ± 0,11

Figura 30

Resultados de modelos principales sobre el set de datos P-UCH.

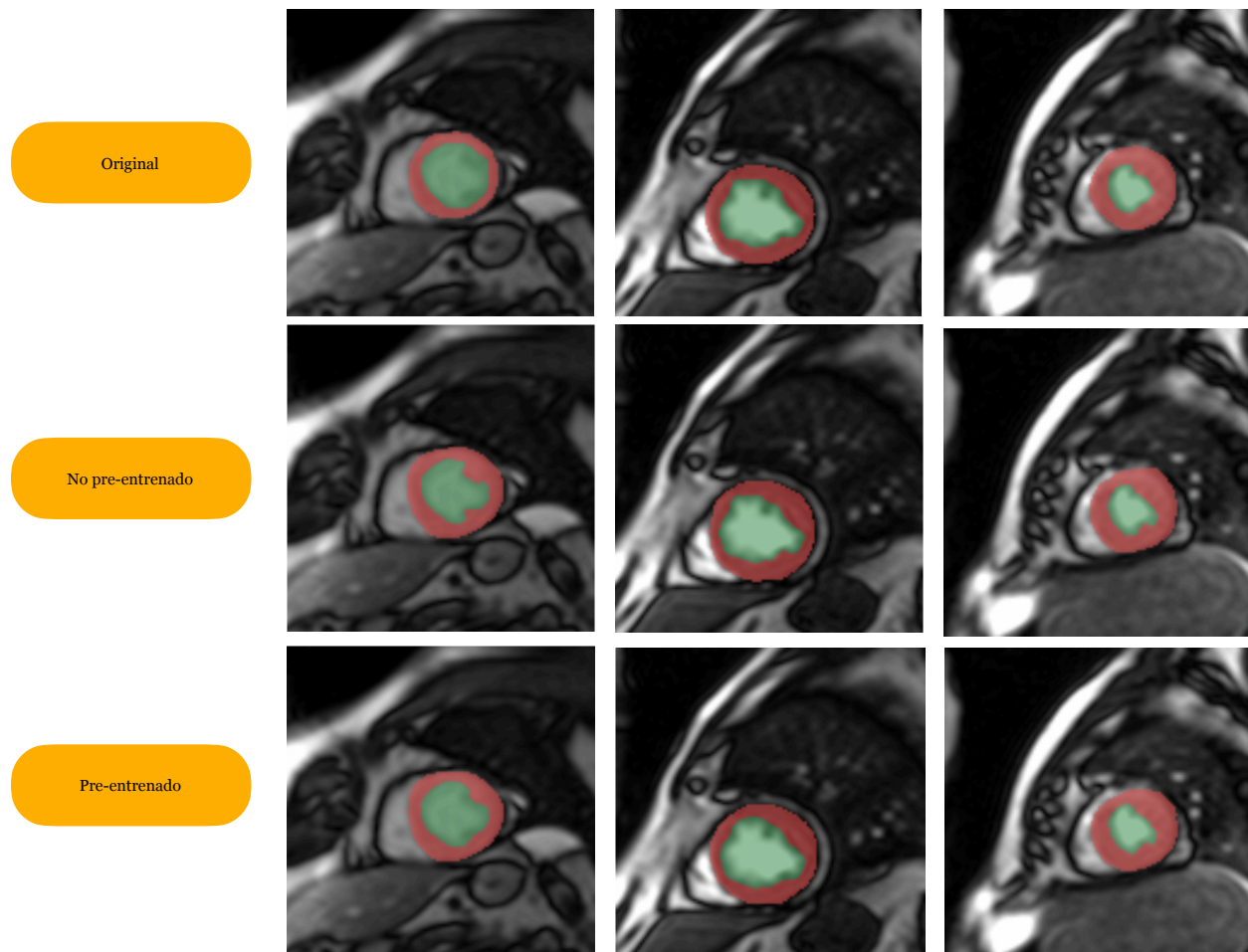


Figura 31

Comparativa resultados de segmentación u-net, con tres imágenes de ejemplo del set de datos de testeo. Primera fila, segmentación manual, segunda fila es una TransUNet no pre-entrenada y tercera fila es la TransUNet pre-entrenada en imágenes de cine.

A continuación, se disponen tableros donde se pueden observar resultados de las segmentaciones, las métricas por cada modelo, los entrenamientos y la validación de los experimentos con pre-entrenamiento y los sin pre-entrenamiento.

- Sin pre-entrenamiento: [reporte 1](#)
- Con pre-entrenamiento: [reporte 2](#)

Resumen

A continuación, se presenta la tabla resumen de los resultados obtenidos de los experimentos descritos anteriormente en la sección de cada modelo utilizado. Los mejores resultados sobre el set de datos P-UCH fueron obtenidos por el modelo que combina la arquitectura U-Net con aprendizaje adversarial profundo y los mejores resultados sobre el set de datos P-GIGA fueron obtenidos por la arquitectura TransUNet en el miocardio y el epicardio, y por la U-Net con aprendizaje adversarial profundo en el endocardio. Por otro lado, al comparar los algoritmos por sus características relevantes para implementación en la Figura 34, se puede ver que la que tiene menos parámetros es la TransUNet pero a la vez es el modelo de mayor tamaño.

Modelo	Pre-entrenamiento	Dice miocardio (%) \pm std (%)	Dice endocardio (%) \pm std (%)	Dice epicardio (%) \pm std (%)	IoU miocardio (%) \pm std (%)	IoU endocardio (%) \pm std (%)	IoU epicardio (%) \pm std (%)
U-Net	✗	82,51 \pm 0,51	92,3 \pm 0,090	93,5 \pm 0,17	83,98 \pm 0,40	92,36 \pm 0,086	93,28 \pm 0,17
	✓	82,48 \pm 0,29	92,18 \pm 0,24	93,53 \pm 0,096	83,94 \pm 0,23	92,24 \pm 0,22	93,3 \pm 0,083
U-Net + Adv. learning	✗	83,98 \pm 0,34	94,07 \pm 0,11	95,55 \pm 0,12	85,58 \pm 0,27	94,16 \pm 0,10	95,32 \pm 0,12
	✓	84,08 \pm 0,22	94,1 \pm 0,14	95,57 \pm 0,11	85,65 \pm 0,17	94,18 \pm 0,13	95,34 \pm 0,11
TransUNet	✗	83,21 \pm 0,59	93,75 \pm 0,20	95,3 \pm 0,21	85,01 \pm 0,44	93,88 \pm 0,17	95,08 \pm 0,20
	✓	83,55 \pm 0,34	93,89 \pm 0,15	95,42 \pm 0,11	85,27 \pm 0,26	93,99 \pm 0,14	95,19 \pm 0,11

Figura 32

Comparativa resultados sobre el set de datos P-UCH.

Modelo	Pre-entrenamiento	Dice miocardio (%) \pm std (%)	Dice endocardio (%) \pm std (%)	Dice epicardio (%) \pm std (%)	IoU miocardio (%) \pm std (%)	IoU endocardio (%) \pm std (%)	IoU epicardio (%) \pm std (%)
U-Net	✗	61,69 \pm 8,96	83,17 \pm 7,23	78,61 \pm 7,67	71,04 \pm 4,73	85,47 \pm 4,90	81,16 \pm 5,22
	✓	67,33 \pm 4,23	85,37 \pm 6,42	82,42 \pm 4,08	74 \pm 2,60	87,13 \pm 4,41	83,77 \pm 3,12
U-Net + Adv. learning	✗	65,73 \pm 4,68	86,21 \pm 4,16	82,29 \pm 3,87	73,25 \pm 2,68	87,67 \pm 3	83,74 \pm 2,89
	✓	69,16 \pm 5,57	87,91 \pm 1,79	84,09 \pm 3,54	75,27 \pm 3,23	88,95 \pm 1,41	85,16 \pm 2,69
TransUNet	✗	64,63 \pm 5,04	83,02 \pm 4,13	80,09 \pm 4,21	72,61 \pm 2,84	85,49 \pm 2,78	82,23 \pm 2,88
	✓	69,38 \pm 3,62	87,55 \pm 1,25	84,19 \pm 2,13	75,42 \pm 2,13	88,72 \pm 0,97	85,2 \pm 1,61

Figura 33
Comparativa resultados sobre el set de datos P-GIGA.

Modelo	Cantidad de parámetros	Tamaño del modelo (MB)
U-Net	310.366.111	118,5
U-Net + Adv. learning	310.366.111	118,5
TransUNet	56.804.576	216,8

Figura 34
Comparativa de características de los modelos.

En la Figura 35, se muestran algunos resultados de la segmentación realizada por el mejor modelo entrenado de cada arquitectura. Se aplican sobre todos los fotogramas de un mismo paciente seleccionado arbitrariamente por su dificultad en las segmentaciones y, de esta manera, observar las diferencias de los resultados de los modelos en imágenes complejas.

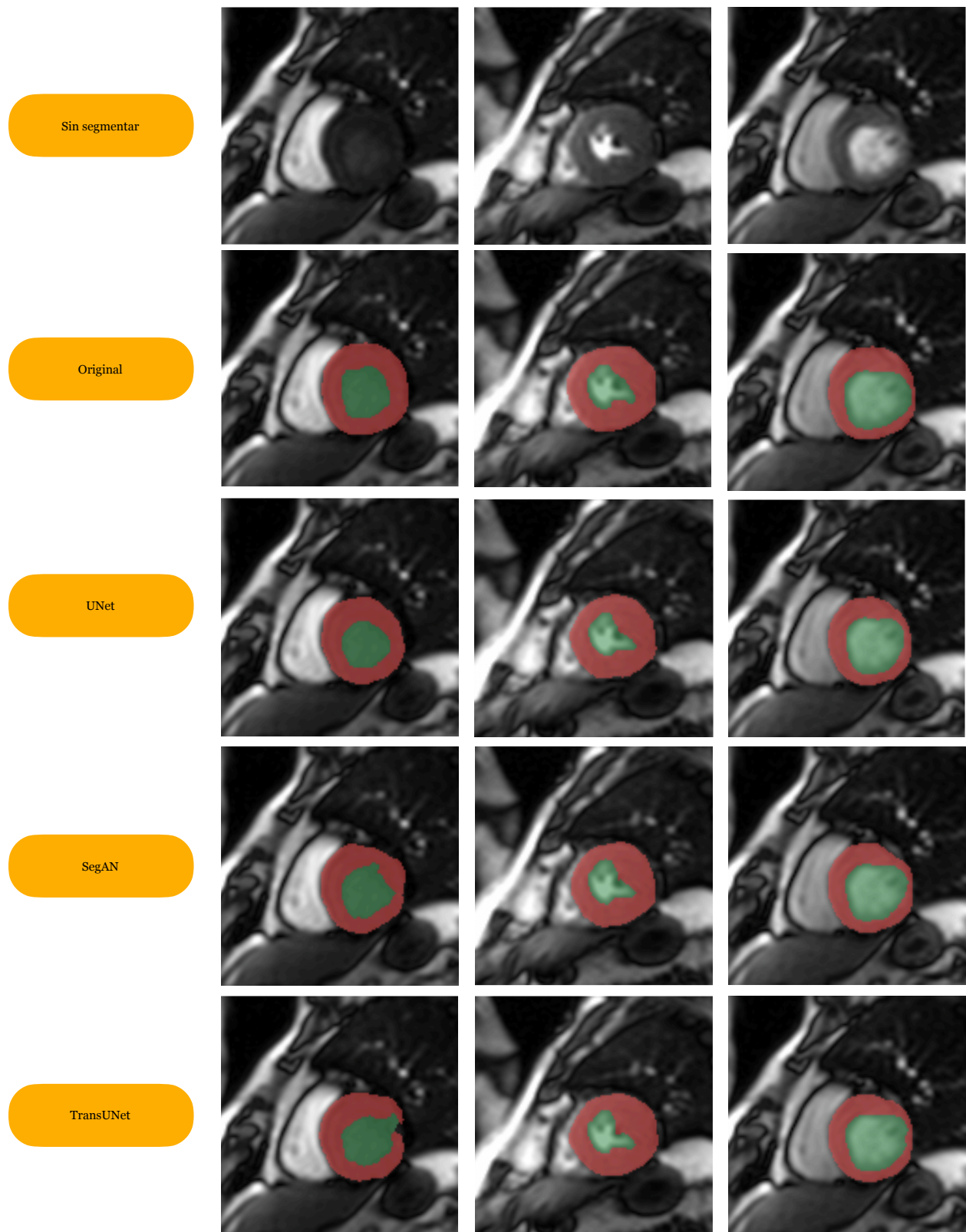


Figura 35

Comparativa resultados de segmentación en imágenes de ejemplo con todos los modelos.

Conclusiones

Contribuciones

En base a los resultados obtenidos en las secciones previas, se obtienen como contribuciones los tres siguientes aspectos que permiten cumplir con el objetivo general y los objetivos específicos planteados al comienzo de la investigación.

Primero, se propone una metodología para la revisión del etiquetado de las imágenes de perfusión cardíaca de resonancia magnética cardiovascular, que entrega recomendaciones para segmentar el borde del endocardio y del epicardio. De esta manera, se puede realizar una segmentación que no limite los resultados de los modelos y disminuya el sesgo presente en los variados criterios utilizados por diferentes especialistas. La metodología planteada se basa en información del artículo de [Bremerich et al. \(2013\)](#) y la exhaustiva revisión de la base de datos de [Kellman y Wollny \(2014\)](#) que fue etiquetada por especialistas.

Segundo, se introduce un entrenamiento con aprendizaje adversarial profundo y una función de pérdida adversarial para mejorar los resultados de los algoritmos de segmentación de imágenes como la U-Net. De esta manera, se fuerza a los modelos a aprender características más finas de las imágenes, debido a que la función de pérdida comprende características básicas, intermedias y de alto nivel. Además, los dos modelos entrenados de forma adversaria se fuerzan uno al otro a realizar mejor su tarea, que en el caso de la red segmentadora es el etiquetado del endocardio y miocardio.

Tercero, como consecuencia de las contribuciones anteriormente descritas, se obtiene un modelo entrenado para la segmentación de la perfusión cardíaca en imágenes de resonancia magnética, con las métricas presentadas en la sección anterior. Esto permite reducir el tiempo de trabajo en la segmentación de imágenes de un paciente desde horas a minutos. Además, contribuye a mejorar la exactitud con la que se detectan las enfermedades cardiovasculares

debido a que permite emplear con efectividad el análisis cuantitativo de las imágenes, en desmedro del análisis únicamente visual utilizado en la práctica clínica habitual.

Discusión

Al comparar los resultados de los modelos, se puede observar que el algoritmo que utiliza el aprendizaje adversarial profundo obtuvo mejores resultados que las otras arquitecturas planteadas. Una posible explicación es que la función de pérdida adversarial utilizada logra estabilizar de mejor manera las segmentaciones realizadas por los modelos. Esto se puede observar en que la desviación estándar de las métricas obtenidas en la mayoría de las veces fue comparativamente menor a las de los otros modelos.

Una explicación de las métricas obtenidas en el miocardio, está relacionada a que dicha etiqueta no fue corregida y, por lo tanto, estandarizada de la manera adecuada para que los modelos puedan aprender una mejor forma de segmentación de la estructura. De esta manera, las métricas están todavía limitadas por la gran variabilidad que se puede observar en la forma de etiquetar dicha clase. En contraste, las métricas observadas en el endocardio tienen valores muy altos y se obtienen segmentaciones bastante precisas.

Por otro lado, en las pruebas de validación externa del algoritmo se tuvo consistencia con lo descrito en el párrafo anterior. Primero, en la etiqueta relacionada al miocardio hubo una caída desde un 84,08% a 69,38% en los modelos con mejores resultados. Lo anterior, se puede explicar debido a la diferencia existente entre las imágenes del set de datos PGIGA y PUCH, que provienen de diferentes dispositivos de captura, por lo que se podría deducir que en dicha clase al modelo le costó generalizar debido a una caída de 14,7% en la métrica Dice. Segundo, en la etiqueta relacionada el endocardio hubo una caída menor de 94,10% a 87,91%. Esto resulta en un buen indicativo del trabajo realizado en la revisión de dicha etiqueta, al tener solo una reducción de un 6,19% que es menos de la mitad que la etiqueta que no fue corregida.

En base a los resultados obtenidos, tanto en el miocardio como en el endocardio, se propone combinar la segmentación automática del algoritmo entrenado, con la supervisión y

correcciones menores de un radiólogo. De esta manera, se reduce drásticamente la carga de trabajo de estos profesionales con una alta calidad en las segmentaciones. Además, con dicha forma de implementación se logran reducir los riesgos de gestión del cambio, existentes en las aplicaciones tecnológicas en la medicina. Lo anterior, se plantea tomando en consideración lo realizado por [Leibig et al. \(2022\)](#) con su implementación de un sistema de detección de cáncer de mamas para asistir a los radiólogos, produciendo resultados más precisos que cuando trabaja cada uno por si solo y cumpliendo con aliviar la carga de trabajo de los profesionales.

Trabajos futuros

El principal trabajo futuro está en la corrección de la etiqueta relacionada al miocardio que implica delimitar el borde del epicardio. Al analizar las métricas obtenidas y los entrenamientos, es posible observar un estancamiento de los resultados, a diferencia de lo que se observa en el endocardio, donde las métricas superan el 94% en *Dice Score* e *IoU score*. De esta manera, se espera estandarizar la clase relacionada al miocardio para que los modelos logren comprender adecuadamente el proceso de segmentación con un mismo criterio en los datos etiquetados.

Por otro lado, es importante destacar la importancia de disponer el modelo entrenado en un *software* que permita a los especialistas del Hospital Clínico de la Universidad de Chile poder realizar el proceso de segmentación de forma automática. De esta manera, pueden enfocar su trabajo en correcciones menores y el adecuado análisis de las imágenes para la detección de enfermedades cardiovasculares.

Bibliografía

- Adeli, E., Chen, J., Lu, Y., Luo, X., Lu, L., Wang, Y., Yu, Q., Yuille, A., Zhou, Y. (2021). TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. ResearchGate. <https://doi.org/10.48550/arXiv.2102.04306>.
- Akbari, M., Karimi, N., Mohrekeh, M., Nasr-Esfahani, E., Nasr-Esfahani, M., Najarian, K., Samavi, S., Soroushmehr, S. (2018). Left Ventricle Segmentation in Cardiac MR Images Using Fully Convolutional Network. IEEE Xplore. <https://doi.org/10.1109/EMBC.2018.8512536>.
- Arai, A., Hsu, L., Jacobs, M., Sandfort, V. (2020). Reliable segmentation of 2D cardiac magnetic resonance perfusion image sequences using time as the 3rd dimension. European Radiology. <https://doi.org/10.1007/s00330-020-07474-5>.
- Bernard, O., Lalande, A., Zotti, C., Cervenansky, F. (2018). Deep Learning Techniques for Automatic MRI Cardiac Multi-structures Segmentation and Diagnosis: Is the Problem Solved?. IEEE Transactions on Medical Imaging. <https://doi.org/10.1109/TMI.2018.2837502>.
- Breeuwer, M., Chiribiri, A., Lee, J., Sammut, E., Scannell, C., Veta, M., Villa, A. (2019). Deep-Learning-Based Preprocessing for Quantitative Myocardial Perfusion MRI. Journal of Magnetic Resonance Imaging. <https://doi.org/10.1002/jmri.26983>.
- Bremerich, J., Bluemke, D., Flamm, S., Friedrich, M., Fogel, M., Knobelsdorff-Brenkenhoff, F., Krammer, C., Kim, R., Nagel, E., Pennell, D., Plein, S., Schulz-Menger, J. (2013). Standardized image interpretation and post processing in cardiovascular magnetic resonance: society for cardiovascular magnetic resonance board of trustees task force on standardized post processing. Journal of Cardiovascular Magnetic Resonance. <https://doi.org/10.1186/1532-429X-15-35>.

- Brown, L., Davies, R., Fontana, M., Kellman, P., Knott, K., Kotecha, T., Moon, J., Plein, S., Xue, H. (2020). Automated Inline Analysis of Myocardial Perfusion MRI with Deep Learning. *Radiology: Artificial Intelligence*. <https://doi.org/10.1148/ryai.2020200009>.
- Choe, Y., Kim, K., Kim, Y. (2019). Automatic myocardial segmentation in dynamic contrast enhanced perfusion MRI using Monte Carlo dropout in an encoder-decoder convolutional neural network. *Computer Methods and Programs in Biomedicine*. <https://doi.org/10.1016/j.cmpb.2019.105150>.
- Dumoulin, V., Visin, F. (2018). A guide to convolution arithmetic for deep learning. *ResearchGate*. <https://doi.org/10.48550/arXiv.1603.07285>.
- Fei, H., Huang, M., Li, M., Ma, D., Shi, Y., Wang, Y., Xie, Q., Xu, X., Zeng, D. (2021). A deep learning approach with temporal consistency for automatic myocardial segmentation of quantitative myocardial contrast echocardiography. *Springer Nature*. <https://doi.org/10.1007/s10554-021-02181-8>.
- García, G., Jimenez, A., Ramos, C., Reyes, E., Tapia, N. (2021). Deep Transfer Learning for Myocardial Perfusion Segmentation in Cardiac MRI. *IEEE transactions on medical imaging*.
- Goodfellow, I., Bengio, Y., Courville, A., Mirza, M., Ozair, S., Pouget-Abadie, J., Warde-Farley, D. y Xu, B. (2014). Generative Adversarial Nets. *NeurIPS*. <https://doi.org/10.48550/arXiv.1406.2661>.
- Hanin, B., Roberts, D., Yaida, S. (2022). *The Principles of Deep Learning Theory*. Cambridge University Press. <https://doi.org/10.48550/arXiv.2106.10165>.
- Harel, J., Koch, C., Perona, P. (2006). Graph-Based Visual Saliency. *NeuralIPS*.
- Huang, X., Rodney Long, L., Xu, T., Xue, Y. y Zhang, H. (2017). SegAN: Adversarial Network with Multi-scale L_1 Loss for Medical Image Segmentation. *Springer link*. <https://doi.org/10.1007/s12021-018-9377-x>.
- Jacobs, M., Arai, A., Benovoy, M., Berry, C., Chang, L., Corcoran, D., Hsu, L. (2021). Automated Segmental Analysis of Fully Quantitative Myocardial Blood Flow Maps by First-Pass

- Perfusion Cardiovascular Magnetic Resonance. IEEE Access. <https://doi.org/10.1109/ACCESS.2021.3070320>.
- Jerosch-Herold, M. (2010). Quantification of myocardial perfusion by cardiovascular magnetic resonance. Journal of Cardiovascular Magnetic Resonance. <https://doi.org/10.1186/1532-429X-12-57>.
- Kramer, F., Muller, D., Soto, I. (2022). Towards a Guideline for Evaluation Metrics in Medical Image Segmentation. <https://doi.org/10.48550/arXiv.2202.05273>.
- Kelle, S., Nagel, E. (2008). The Role of Magnetic Resonance Imaging in the Detection of Coronary Artery Disease. US Cardiology Volume 5 Issue 1. <https://doi.org/10.15420/usc.2008.5.1.31>.
- Kellman, P., Wollny, G. (2014). Free breathing myocardial perfusion data sets for performance analysis of motion compensation algorithms. ResearchGate. <https://doi.org/10.1186/2047-217X-3-23>.
- Leibig, C., Brehmer, M., Bunk, S., Byng, D., Pinker, K., Umutlu, L. (2022). Combining the strengths of radiologists and AI for breast cancer screening: a retrospective analysis. The Lancet Digital Health. [https://doi.org/10.1016/S2589-7500\(22\)00070-X](https://doi.org/10.1016/S2589-7500(22)00070-X).
- Ronneberger, O., Brox, T., Fischer, P. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. Springer International Publishing Switzerland. <https://doi.org/10.48550/arXiv.1505.04597>.
- Texas Heart Institute. (2022). Enfermedad arterial coronaria. <https://www.texasheart.org/heart-health/heart-information-center/topics/enfermedad-arterial-coronaria/>.
- Vallejo, E. (2009). Enfermedad arterial coronaria o cardiopatía isquémica: dos entidades distintas con diferentes procedimientos diagnósticos. Servicio de Cardiología Nuclear y Tomografía Cardíaca. Instituto Nacional de Cardiología Ignacio Chávez.
- Vaswani, A., Gomez, A., Jones, L., Kaiser, L., Parmar, N., Polosukhin, I., Shazeer, N., Uszkoreit, J. (2017). Attention Is All You Need. NeuralIPS. <https://doi.org/10.48550/arXiv.1706.03762>.

Wang, J., Wang, T., Zhang, Y., Zhao, J. (2021). A Myocardial Segmentation Method Based on Adversarial Learning. Hindi: BioMed Research International. <https://doi.org/10.1155/2021/6618918>.

World Health Organization. (2022). Organización World Health Statistics 2022: monitoring health for the SDGs, sustainable development goals. Geneva: World Health Organization; 2022. License: CC BY-NC-SA 3.0 IGO.

Anexo

Experimentos

U-Net

Debido a que uno de los objetivos de la tesis es obtener los mejores resultados posibles y tener una capacidad alta de generalización del modelo, se realizan diferentes experimentos para poder definir algunos hiperparámetros relevantes como la cantidad de features iniciales, la tasa de aprendizaje inicial, el factor de decaimiento de la tasa de aprendizaje, el tamaño del lote y el uso del decaimiento del peso como método de regularización.

En la fase 1, se enfoca en analizar la cantidad de features iniciales a utilizar, dejando el resto de hiperparámetros en forma estándar sin modificar, y de esa forma se llega al comparar las métricas que se converge a mejores resultados utilizando 64 features. En la fase 2, se analiza la tasa de aprendizaje, que si bien se llega a mejores resultados con las tasa de aprendizaje más alta, el entrenamiento es más estable con las otras tazas de aprendizaje, por lo que se decide utilizar 0.0001. En la fase 3, se analiza el valor del decaimiento del peso, que al comparar dos opciones que son las más utilizadas por la comunidad, se llega a que el entrenamiento es más estable con 0.001 y se obtienen resultados similares.

Fase	Features iniciales	Learning rate	Batch	Weight decay	Dice endocardio	Dice miocardio	Dice epicardio
1	32	0.0001	36	-	0,9203	0,8177	0,9316
1	32	0.0001	36	-	0,9192	0,8171	0,9338
1	32	0.0001	36	-	0,9218	0,8195	0,9325
Promedio					0,9203	0,8177	0,9316
1	64	0.0001	36	-	0,9225	0,8246	0,9349

Fase	Features iniciales	Learning rate	Batch	Weight decay	Dice endocardio	Dice miocardio	Dice epicardio
1	64	0.0001	36	-	0,9235	0,8241	0,9351
1	64	0.0001	36	-	0,9242	0,8251	0,9359
Promedio					0,9242	0,8251	0,9359
2	64	0.001	36	-	0,9256	0,8300	0,9370
2	64	0.001	36	-	0,9233	0,8296	0,9371
2	64	0.001	36	-	0,9195	0,8252	0,9365
Promedio					0,9228	0,8283	0,9369
2	64	0.0001	36	-	0,9222	0,8210	0,9352
2	64	0.0001	36	-	0,9245	0,8253	0,9355
2	64	0.0001	36	-	0,9232	0,8229	0,9346
Promedio					0,9233	0,8241	0,9351
2	64	0.00001	36	-	0,9137	0,8062	0,9302
2	64	0.00001	36	-	0,9144	0,8055	0,9315
2	64	0.00001	36	-	0,9142	0,8080	0,9318
Promedio					0,9141	0,8066	0,9312
3	64	0.0001	36	0.01	0,9198	0,8203	0,9330
3	64	0.0001	36	0.01	0,9241	0,8282	0,9354
3	64	0.0001	36	0.01	0,9168	0,8185	0,9320
Promedio					0,9202	0,8223	0,9335
3	64	0.0001	36	0.001	0,9215	0,8246	0,9347
3	64	0.0001	36	0.001	0,9196	0,8208	0,9326
3	64	0.0001	36	0.001	0,9232	0,8203	0,9330

Fase	Features iniciales	Learning rate	Batch	Weight decay	Dice endocardio	Dice miocardio	Dice epicardio
Promedio					0,9214	0,8219	0,9334

Figura 36
Búsqueda de hiperparámetros específicos de la arquitectura.

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Experimento 1	0,8247	0,9223	0,9358	0,8395	0,9229	0,9336
Experimento 2	0,8304	0,9241	0,9371	0,8439	0,9246	0,9349
Experimento 3	0,8287	0,9227	0,9359	0,8428	0,9233	0,9337
Experimento 4	0,8161	0,9240	0,9318	0,8328	0,9244	0,9297
Experimento 5	0,8302	0,9236	0,9361	0,8437	0,9242	0,9339
Experimento 6	0,8228	0,9233	0,9346	0,8380	0,9238	0,9324
Experimento 7	0,8167	0,9225	0,9322	0,8331	0,9230	0,9300
Experimento 8	0,8269	0,9232	0,9360	0,8416	0,9237	0,9338
Experimento 9	0,8277	0,9238	0,9357	0,8418	0,9243	0,9335
Experimento 10	0,8274	0,9212	0,9355	0,8417	0,9218	0,9333
Promedio	0,8252	0,9231	0,9351	0,8399	0,9236	0,9329
Desviación estándar	0,0051567	0,0009044	0,0017321	0,0040690	0,0008641	0,0017113

Figura 37
Resultados obtenidos de 10 experimentos sin pre-entrenamiento sobre el set de datos PUCH.

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Experimento 1	0,8266	0,9248	0,9347	0,8409	0,9251	0,9325
Experimento 2	0,8219	0,9194	0,9348	0,8372	0,9203	0,9326
Experimento 3	0,8248	0,9230	0,9355	0,8394	0,9235	0,9333
Experimento 4	0,8258	0,9240	0,9349	0,8402	0,9245	0,9327
Experimento 5	0,8209	0,9176	0,9336	0,8366	0,9184	0,9314
Experimento 6	0,8306	0,9239	0,9366	0,8440	0,9244	0,9344
Experimento 7	0,8217	0,9211	0,9351	0,8371	0,9217	0,9329
Experimento 8	0,8267	0,9243	0,9351	0,8408	0,9247	0,9328
Experimento 9	0,8230	0,9206	0,9361	0,8379	0,9213	0,9339
Experimento 10	0,8264	0,9202	0,9368	0,8408	0,9210	0,9336
Promedio	0,8248	0,9219	0,9353	0,8395	0,9225	0,9330
Desviación estándar	0,00299080	0,00244379	0,00096356	0,00231202	0,00226542	0,00083858

Figura 38

Resultados obtenidos de 10 experimentos con pre-entrenamiento sobre el set de datos PUCH.

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Experimento 1	0,6395	0,8924	0,8232	0,7204	0,8995	0,8359
Experimento 2	0,6482	0,8375	0,8130	0,7259	0,8564	0,8298
Experimento 3	0,7214	0,8646	0,8672	0,7686	0,8767	0,8717
Experimento 4	0,5049	0,7721	0,6926	0,6498	0,8099	0,7458

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Experimento 5	0,7198	0,8431	0,8487	0,7697	0,8615	0,8568
Experimento 6	0,6163	0,8694	0,7958	0,7069	0,8806	0,8149
Experimento 7	0,6060	0,8463	0,7758	0,7026	0,8632	0,8017
Experimento 8	0,6267	0,8620	0,8055	0,7126	0,8752	0,8237
Experimento 9	0,6561	0,8808	0,8250	0,7307	0,8899	0,8374
Experimento 10	0,4298	0,6483	0,6142	0,6166	0,7336	0,6987
Promedio	0,6169	0,8317	0,7861	0,7104	0,8547	0,8116
Desviación estándar	0,0896297	0,0723276	0,0767229	0,0473812	0,0490065	0,0522534

Figura 39

Resultados obtenidos de 10 experimentos sin pre-entrenamiento, y prueba sobre el set de datos PGIGA.

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Experimento 1	0,6841	0,8775	0,8415	0,7472	0,8888	0,8501
Experimento 2	0,7617	0,9000	0,8870	0,7954	0,9057	0,8880
Experimento 3	0,6520	0,8929	0,8300	0,7265	0,8998	0,8407
Experimento 4	0,6414	0,8737	0,8095	0,7203	0,8844	0,8258
Experimento 5	0,7255	0,8797	0,8716	0,7711	0,8888	0,8750
Experimento 6	0,6823	0,8513	0,8260	0,7448	0,8671	0,8383
Experimento 7	0,6194	0,8444	0,7946	0,7079	0,8617	0,8143
Experimento 8	0,6544	0,9013	0,8255	0,7290	0,9069	0,8384
Experimento 9	0,6625	0,8328	0,8180	0,7333	0,8526	0,8322
Experimento 10	0,6494	0,6833	0,7383	0,7245	0,7572	0,7745

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Promedio	0,6733	0,8537	0,8242	0,7400	0,8713	0,8377
Desviación estándar	0,0423722	0,0642510	0,0408765	0,0260728	0,0441278	0,0312670

Figura 40

Resultados obtenidos de 10 experimentos con pre-entrenamiento, y prueba sobre el set de datos PGIGA.

U-Net + Aprendizaje Adversarial Profundo

A continuación, se muestra la experimentación realizada para definir el factor α de la función de pérdida utilizada para entrenar la U-Net con aprendizaje adversarial profundo.

Experimento	Alpha	Dice miocardio	Dice endocardio	Dice epicardio
1	0.05	0,8377	0,9400	0,9553
2	0.05	0,8402	0,9413	0,9551
Promedio		0,83895	0,94065	0,9552
1	0.10	0,8395	0,9390	0,9576
2	0.10	0,8408	0,9404	0,9560
Promedio		0,84015	0,9397	0,9568
1	0.15	0,8398	0,9406	0,9561
2	0.15	0,8374	0,9399	0,9548
Promedio		0,8386	0,9403	0,9555
1	0.2	0,8401	0,9402	0,9551
2	0.2	0,8382	0,9414	0,9557
Promedio		0,8392	0,9408	0,9554

Figura 41

Comparativa resultados obtenidos con SegAN y diferentes alpha de entrada.

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Experimento 1	0,8388	0,9410	0,9558	0,8550	0,9418	0,9535
Experimento 2	0,8384	0,9418	0,9531	0,8545	0,9426	0,9509
Experimento 3	0,8426	0,9414	0,9572	0,8581	0,9422	0,9549
Experimento 4	0,8349	0,9400	0,9552	0,8520	0,9409	0,9529
Experimento 5	0,8355	0,9410	0,9539	0,8523	0,9418	0,9517
Experimento 6	0,8422	0,9390	0,9564	0,8578	0,9401	0,9541
Experimento 7	0,8460	0,9429	0,9569	0,8606	0,9435	0,9546
Experimento 8	0,8408	0,9398	0,9554	0,8565	0,9408	0,9531
Experimento 9	0,8374	0,9402	0,9550	0,8540	0,9411	0,9527
Experimento 10	0,8414	0,9399	0,9560	0,8570	0,9410	0,9537
Promedio	0,8398	0,9407	0,9555	0,8558	0,9416	0,9532
Desviación estándar	0,0034451	0,0011450	0,0012767	0,0027186	0,0010020	0,0012423

Figura 42

Resultados obtenidos de 10 experimentos sin pre-entrenamiento en el set de datos PUCH.

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Experimento 1	0,8414	0,9412	0,9535	0,8569	0,9419	0,9513
Experimento 2	0,8428	0,9424	0,9552	0,8580	0,9431	0,9529
Experimento 3	0,8406	0,9390	0,9554	0,8565	0,9401	0,9531
Experimento 4	0,8390	0,9396	0,9563	0,8552	0,9404	0,9541
Experimento 5	0,8382	0,9394	0,9569	0,8545	0,9403	0,9547
Experimento 6	0,8393	0,9416	0,9555	0,8552	0,9423	0,9532
Experimento 7	0,8378	0,9402	0,9545	0,8542	0,9411	0,9522
Experimento 8	0,8412	0,9433	0,9562	0,8569	0,9438	0,9540
Experimento 9	0,8429	0,9417	0,9560	0,8581	0,9425	0,9537
Experimento 10	0,8446	0,9412	0,9572	0,8595	0,9421	0,9549
Promedio	0,8408	0,9410	0,9557	0,8565	0,9418	0,9534
Desviación estándar	0,0022	0,0014	0,0011	0,0017	0,0013	0,0011

Figura 43

Resultados obtenidos de 10 experimentos con pre-entrenamiento en el set de datos PUCH.

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Experimento 1	0,6879	0,8837	0,8475	0,7488	0,8925	0,8554
Experimento 2	0,6596	0,8790	0,8254	0,7325	0,8890	0,8388
Experimento 3	0,6650	0,8685	0,8245	0,7367	0,8809	0,8382
Experimento 4	0,6185	0,8757	0,8061	0,7114	0,8864	0,8239
Experimento 5	0,6346	0,8721	0,8176	0,7179	0,8836	0,8323
Experimento 6	0,7359	0,8621	0,8775	0,7782	0,8755	0,8801
Experimento 7	0,6343	0,8703	0,8170	0,7198	0,8823	0,8319
Experimento 8	0,7106	0,8769	0,8555	0,7644	0,8876	0,8629
Experimento 9	0,6537	0,8874	0,8274	0,7299	0,8960	0,8396
Experimento 10	0,5727	0,7455	0,7308	0,6851	0,7929	0,7706
Promedio	0,6573	0,8621	0,8229	0,7325	0,8767	0,8374
Desviación estándar	0,0468	0,0416	0,0387	0,0268	0,0300	0,0289

Figura 44

Resultados obtenidos de 10 experimentos sin pre-entrenamiento, y prueba sobre el set de datos PGIGA.

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Experimento 1	0,7342	0,8960	0,8693	0,7776	0,9027	0,8731
Experimento 2	0,6858	0,8651	0,8338	0,7473	0,8780	0,8449
Experimento 3	0,7405	0,8960	0,8720	0,7814	0,9033	0,8751
Experimento 4	0,6397	0,8510	0,8069	0,7228	0,8680	0,8254
Experimento 5	0,5739	0,8662	0,7709	0,6860	0,8793	0,7985
Experimento 6	0,7273	0,9002	0,8663	0,7745	0,9063	0,8712
Experimento 7	0,7320	0,8873	0,8695	0,7765	0,8959	0,8734
Experimento 8	0,7234	0,8879	0,8633	0,7717	0,8962	0,8691
Experimento 9	0,6408	0,8557	0,8047	0,7215	0,8710	0,8243
Experimento 10	0,7179	0,8851	0,8527	0,7681	0,8939	0,8605
Promedio	0,6916	0,8791	0,8409	0,7527	0,8895	0,8516
Desviación estándar	0,0557541	0,0179402	0,0354255	0,0323677	0,0141142	0,0269888

Figura 45

Resultados obtenidos de 10 experimentos con pre-entrenamiento, y prueba sobre el set de datos PGIGA.

TransUNet

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Experimento 1	0,8385	0,9391	0,9538	0,8549	0,9402	0,9514
Experimento 2	0,8344	0,9380	0,9517	0,8517	0,9392	0,9494
Experimento 3	0,8270	0,9396	0,9522	0,8462	0,9405	0,9500

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Experimento 4	0,8321	0,9389	0,9543	0,8502	0,9400	0,9520
Experimento 5	0,8184	0,9359	0,9481	0,8400	0,9375	0,9463
Experimento 6	0,8319	0,9380	0,9535	0,8499	0,9391	0,9513
Experimento 7	0,8297	0,9370	0,9536	0,8482	0,9386	0,9515
Experimento 8	0,8355	0,9374	0,9551	0,8527	0,9384	0,9529
Experimento 9	0,8377	0,9386	0,9557	0,8544	0,9397	0,9535
Experimento 10	0,8360	0,9324	0,9520	0,8531	0,9345	0,9499
Promedio	0,8321	0,9375	0,9530	0,8501	0,9388	0,9508
Desviación estándar	0,0059989	0,0020941	0,0021597	0,0044811	0,0017550	0,0020520

Figura 46

Resultados obtenidos de 10 experimentos sin pre-entrenamiento sobre el set de datos PUCH.

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Experimento 1	0,8334	0,9376	0,9527	0,8513	0,9387	0,9505
Experimento 2	0,8317	0,9365	0,9535	0,8498	0,9377	0,9512
Experimento 3	0,8363	0,9377	0,9539	0,8533	0,9389	0,9516
Experimento 4	0,8343	0,9392	0,9550	0,8516	0,9402	0,9527
Experimento 5	0,8343	0,9402	0,9550	0,8515	0,9411	0,9528
Experimento 6	0,8311	0,9388	0,9529	0,8494	0,9398	0,9506
Experimento 7	0,8381	0,9380	0,9539	0,8549	0,9391	0,9516
Experimento 8	0,8415	0,9414	0,9563	0,8572	0,9423	0,9541

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Experimento 9	0,8399	0,9411	0,9553	0,8561	0,9420	0,9531
Experimento 10	0,8344	0,9386	0,9533	0,8517	0,9396	0,9510
Promedio	0,8355	0,9389	0,9542	0,8527	0,9399	0,9519
Desviación estándar	0,0034166	0,0015871	0,0011698	0,0026229	0,0014796	0,0011952

Figura 47

Resultados obtenidos de 10 experimentos con pre-entrenamiento sobre el set de datos PUCH.

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Experimento 1	0,6792	0,8387	0,8228	0,7439	0,8585	0,8354
Experimento 2	0,7186	0,8680	0,8554	0,7690	0,8832	0,8630
Experimento 3	0,6131	0,8350	0,7840	0,7063	0,8561	0,8079
Experimento 4	0,7150	0,8428	0,8481	0,7652	0,8619	0,8567
Experimento 5	0,6160	0,8036	0,7728	0,7088	0,8354	0,8024
Experimento 6	0,6875	0,8628	0,8364	0,7485	0,8783	0,8463
Experimento 7	0,5847	0,8505	0,7755	0,6913	0,8683	0,8035
Experimento 8	0,6226	0,7899	0,7700	0,7119	0,8258	0,7989
Experimento 9	0,6407	0,8709	0,8204	0,7210	0,8838	0,8343
Experimento 10	0,5851	0,7395	0,7236	0,6953	0,7974	0,7748
Promedio	0,6463	0,8302	0,8009	0,7261	0,8549	0,8223
Desviación estándar	0,0504207	0,0413284	0,0421224	0,0284145	0,0278087	0,0288523

Figura 48

Resultados obtenidos de 10 experimentos sin pre-entrenamiento, y prueba sobre el set de datos PGIGA.

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Experimento 1	0,6558	0,8556	0,8147	0,7317	0,8721	0,8320
Experimento 2	0,7429	0,8723	0,8651	0,7833	0,8835	0,8693
Experimento 3	0,7063	0,8671	0,8468	0,7595	0,8800	0,8548
Experimento 4	0,6683	0,8918	0,8380	0,7382	0,8997	0,8484
Experimento 5	0,6877	0,8755	0,8395	0,7502	0,8874	0,8500

Métricas	Dice miocardio	Dice endocardio	Dice epicardio	IoU miocardio	IoU endocardio	IoU epicardio
Experimento 6	0,7525	0,8979	0,8777	0,7903	0,9048	0,8804
Experimento 7	0,6856	0,8827	0,8397	0,7502	0,8931	0,8495
Experimento 8	0,7055	0,8739	0,8465	0,7599	0,8860	0,8553
Experimento 9	0,6985	0,8727	0,8472	0,7577	0,8855	0,8564
Experimento 10	0,6352	0,8650	0,8038	0,7211	0,8794	0,8243
Promedio	0,6938	0,8755	0,8419	0,7542	0,8872	0,8520
Desviación estándar	0,0362160	0,0125465	0,0213708	0,0213768	0,0097727	0,0161053

Figura 49

Resultados obtenidos de 10 experimentos con pre-entrenamiento, y prueba sobre el set de datos PGIGA.