

E.T.S. de Ingeniería Industrial,  
Informática y de Telecomunicación



# SISTEMA DE DETECCIÓN DE TEJIDO INFARTADO EN EL MIOCARDIO SOBRE IMÁGENES DE RESONANCIA MAGNÉTICA

Máster en Ingeniería Biomédica

Trabajo Fin de Máster

**Autor:** *Idoia Hernandez Echarren*

**Tutores:** *Arantxa Villanueva Larre*

*Rafael Cabeza Laguna*

*Pamplona, 27 de Abril de 2017*

**upna**  
Universidad  
Pública de Navarra  
Nafarroako  
Unibertsitate Publikoa



# Resumen

Las imágenes obtenidas mediante la técnica de realce tardío con Gadolinio (RTG) por resonancia magnética cardíaca (RMC) tienen la capacidad de acentuar las cicatrices generadas por un infarto de miocardio. La cuantificación de este tejido dañado requiere una segmentación fiable de las regiones infartadas. El principal objetivo de este trabajo es la implementación de un sistema de detección de las cicatrices de miocardio en imágenes de RTG por RMC. Con ese propósito, partiendo de imágenes marcadas (miocardio y cicatriz) por un especialista, se han extraído un conjunto de características estadísticas de intensidad y textura para cada píxel del miocardio de las imágenes. Finalmente, parte de las características calculadas han sido empleadas en el entrenamiento de un sistema de clasificación SVM, mientras que el resto han sido utilizadas para verificar el correcto funcionamiento del éste.

## Palabras clave:

Cicatriz de miocardio, Tejido infartado, Resonancia Magnética Cardíaca, Realce Tardío con Gadolinio, detección, SVM

## Abstract

Images obtained by late Gadolinium enhancement (LGE) cardiac magnetic resonance (CMR) technique have the ability to accentuate scars generated by myocardial infarction. For the correct quantification of the damaged tissue is necessary a reliable segmentation of the infarcted regions. The main goal of this report is the implementation of a myocardial infarction scars detection system in LGE CMR images. For this purpose, and starting from images marked (myocardium and scar) by a specialist, a set of statistical features of intensity and texture for each myocardium image pixel have been calculated. Finally, some of the features have been used in the training of a SVM classification systems, while the rest have been use to verify the correct operations of this.

## Keywords:

Myocardial scar, infarcted tissues, Cardiac Magnetic Resonance, late Gadolinium enhancement, detection, SVM.

## Acrónimos

Acc	Accuracy, exactitud
FCM	Fuzzy C-mean
FN	Falso Negativo
FP	Falso Positivo
GLCM	matriz de Co-ocurrencia de nivel de gris
GW	Gabor wavelet
IAM	Infarto Agudo de Miocardio
LBP	Local Binary Pattern
LTE	Laws Texture Energy
Q1 y Q3	Primer y tercer cuartil.
RTG por RMC	Realce tardío de Gadolinio por Resonancia Magnética Cardíaca
Sn	Sensibilidad
Sp	Especificidad
SVM	Support Vector Machine
TN	Verdadero negativo
TP	Verdadero positivo

# Tabla de contenido

<b>Capítulo 1: Introducción.....</b>	<b>1</b>
Cicatrices de miocardio .....	1
Estado del arte.....	3
Objetivos .....	5
Estructura del documento .....	6
<b>Capítulo 2: Base de datos .....</b>	<b>7</b>
<b>Capítulo 3: Sistema propuesto .....</b>	<b>9</b>
<b>Método de pre-procesado.....</b>	<b>9</b>
<b>Extracción de características .....</b>	<b>11</b>
Nivel de gris .....	11
Nivel medio de gris .....	11
Desviación estándar .....	11
Coeficiente de curtosis .....	12
Coeficiente de asimetría de Fisher .....	12
Matriz de Co-ocurrencia de Nivel de Gris (GLCM).....	13
Local Binary Pattern (LBP) .....	15
Laws Texture Energy (LTE).....	16
Gabor wavelet (GW).....	18
Distancias .....	19
<b>Selección de características .....</b>	<b>20</b>
Variables .....	20
Prueba de t-test.....	21
Análisis de correlación.....	22
Simplificación final de las características .....	23
Estandarización .....	23
Organización de las características .....	24
<b>Clasificación de cicatrices .....</b>	<b>25</b>
Support Vector Machine (SVM) .....	25
Evaluación del clasificador .....	27
Implementación .....	29

<b>Capítulo 4 Resultados preliminares.....</b>	<b>35</b>
Validación dejando una imagen para testear: .....	35
Validación dejando un volumen para testear: .....	39
<b>Capítulo 5: Optimización y resultados finales .....</b>	<b>43</b>
Eliminación de imágenes inadecuadas .....	43
Post-procesado .....	44
Comparativa de resultados finales con resultados de CVI <sup>42</sup> .....	47
<b>Capítulo 6: Conclusiones y líneas futuras.....</b>	<b>51</b>
Conclusiones.....	51
Líneas futuras de investigación.....	53
<b>Bibliografía .....</b>	<b>55</b>
<b>Anexo I.....</b>	<b>59</b>

## Listado de figuras

Fig 1. Izquierda: Imagen cardíaca obtenida mediante la técnica de RTG por RMC con el epicardio, endocardio y la cicatriz definidos. Derecha: misma imagen aumentada. ....	2
Fig 2. Imágenes de RMC pertenecientes a la base de datos empleada. ....	7
Fig 3. Ejemplo de la interpolación de las marcas proporcionadas por el especialista. ....	10
Fig 4. Ejemplo de la selección de los píxeles pertenecientes al miocardio definido por el especialista. ....	10
Fig 5. Formas y clasificación de las curvas de distribución, según la <i>curtosis</i> . ....	12
Fig 6. Formas de la curva de distribución según la simetría. $M_d$ es la media y $M_o$ representa la moda, el valor que presenta una mayor frecuencia. ....	13
Fig 7. Ejemplo del cálculo de una GLCM en dirección vertical. ....	14
Fig 8. Ejemplo del proceso llevado a cabo para el cálculo de LBP. ....	16
Fig 9. Mascaras de LTE empleadas. ....	17
Fig 10. Filtros de GW empleados. ....	19
Fig 11. Evolución del tejido infartado o la necrosis con el tiempo. ....	20
Fig 12. Representación de características correlacionadas con un coeficiente mayor que 0.8. ....	22
Fig 13. Organización de las características. ....	25
Fig 14. Ejemplo de obtención de posibles hiperplanos de separación en un conjunto separable en dos clases. ....	26
Fig 15. Hiperplano de separación óptimo logrado con el margen asociado máximo en el ejemplo anterior. ....	26
Fig 16. Ejemplo de la transformación y separación realizada por SVM a un grupo de entrada no separable linealmente. ....	27
Fig 17. Ejemplo de validación cruzada. ....	28
Fig 18. Curva de ROC variando la relación entre píxeles de tejido infartado y sanos. ....	30
Fig 19. Ampliación de la curva ROC anterior. ....	31
Fig 20. Valores de Accuracy, sensibilidad y especificidad de todas las opciones estudiadas. ....	32
Fig 21. División de los datos para trabajar con el clasificador. ....	33
Fig 22. Cuatro ejemplos de los resultados obtenidos empleando la validación cruzada dejando una imagen entera para testear. ....	36
Fig 23. Matriz de resultados de la prueba de validación cruzada dejando una imagen entera para testear. ....	37

Fig 24. Box-plot de los resultados obtenidos en la prueba, dejando imágenes enteras para testear. ....	38
Fig 25. Cuatro ejemplos de los resultados obtenidos empleando una evaluación de validación cruzada dejando un estudio entero para testear. ....	40
Fig 26. Matriz de resultados de la prueba de validación cruzada dejando un estudio entero para testear. ....	41
Fig 27. Blox-pot que recoge los resultados obtenidos en la prueba dejando usuarios completos para testear. ....	42
Fig 28. Ejemplo de tejido infartado sin etiquetar. ....	43
Fig 29. Cuatro ejemplos de los resultados obtenidos tras el post-procesado. ....	45
Fig 30. Box-plot comparativo de los resultados preliminares, marcados con 0, y los post-procesados, representados con el 1. ....	46
Fig 31. Comparación cualitativa de resultados. Izquierda: resultados con el clasificador; Derecha: resultados de CVI <sup>42</sup> . ....	48

# Capítulo 1:

## Introducción

---

### Cicatrices de miocardio

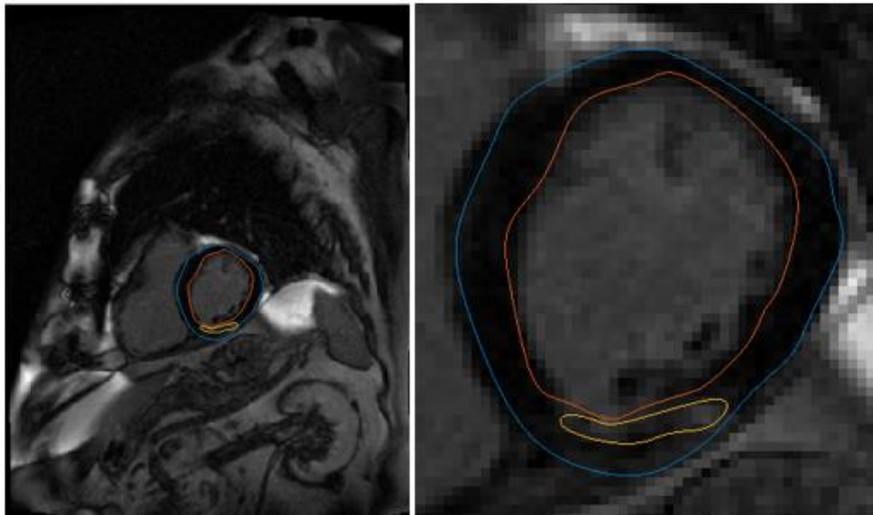
El infarto agudo de miocardio (*IAM*) es la necrosis o la muerte de una porción del músculo cardíaco que se produce cuando se obstruye una de las arterias coronarias y como consecuencia se produce un riego sanguíneo insuficiente [1].

A todos los pacientes que han sufrido un infarto de miocardio les aparecen cicatrices de miocardio o tejidos miocárdicos infartados. Las fibras del músculo miocárdico se encuentran dañadas en todo el tejido infartado, provocando que esta zona no reaccione a la señal eléctrica que se propaga a través del músculo cardíaco o perturbando dicha señal. Estas cicatrices pueden ser causantes de una posterior arritmia cardíaca incapacitante o incluso mortal, por lo que resulta de interés general definir y evaluar cada una de estas cicatrices para así poder realizar un seguimiento adecuado, desarrollar estrategias de intervención, planes de tratamiento y prevenir arritmias futuras en los pacientes.

Las cicatrices generadas tras el infarto de miocardio pueden ser visualizadas y cuantificadas mediante imágenes de resonancia magnética cardíaca (RMC) obtenidas 10-15 minutos después de la administración intravenosa del contraste de Gadolinio. Esta técnica de obtención de imágenes se conoce como realce tardío con Gadolinio (RTG) por resonancia magnética cardíaca [2].

La técnica de RTG por RMC se basa en la captación de gadolinio por parte del miocardio necrosado. El gadolinio es un contraste de elevado coeficiente de partición, siendo más del 60% retenido en el espacio intersticial del miocardio tras su inyección intravenosa. Al mismo tiempo, el espacio intersticial del tejido infartado aumenta, haciendo que la cinética de intercambio del gadolinio sea más lenta en las cicatrices que en el miocardio normal. Ambos efectos resultan en una mayor concentración y mayor persistencia del gadolinio en las cicatrices, permitiendo identificar las zonas con necrosis miocárdica. Además, la resonancia magnética cardíaca consta de una alta resolución espacial y posibilita realizar cortes tomográficos muy finos, permitiendo medir de manera precisa el volumen o la masa del tejido dañado [3].

En la Fig 1 se representa una imagen cardíaca con el epicardio, endocardio y la cicatriz definidos mediante líneas azules, rojas y amarillas, respectivamente.



**Fig 1.** Izquierda: Imagen cardíaca obtenida mediante la técnica de RTG por RMC con el epicardio, endocardio y la cicatriz definidos. Derecha: misma imagen aumentada.

Los modelos experimentales han demostrado un excelente acuerdo entre el tamaño y la forma de las áreas del infarto de miocardio obtenido mediante la técnica de RTG por RMC y por histopatología [4].

Hasta ahora, el método más común para la visualización y definición de las cicatrices de miocardio era realizado de manera manual por un especialista entrenado, teniendo que analizar cada uno de los cortes y definir en todos ellos los tejidos dañados.

Debido a la importancia de la detección de las cicatrices y a la gran carga de trabajo que supone en la actualidad esta labor para los especialistas, se entiende la necesidad de generar un

software que sea capaz de analizar el miocardio y determinar la presencia de tejido infartado en imágenes obtenidas con la técnica de RTG por RMC. Dicha herramienta puede entenderse como una técnica de CAD (*Computer Aided Diagnosis*), un procedimiento médico que ayude a los especialistas en la interpretación de los contenidos multimedia obtenidos en las pruebas a las que se someten los pacientes, en este caso a la técnica de RTG por RMC.

## Estado del arte

A continuación se presentará un resumen de los programas de detección de tejidos infartados publicados hasta la fecha.

La mayoría de los modelos desarrollados se basan en la umbralización de la intensidad de los píxeles pertenecientes al miocardio, ya que gracias a la técnica de RTG por RMC los píxeles perteneciente a los tejidos dañados son notoriamente de mayor intensidad.

Entre los sistemas desarrollados basados en la umbralización de la intensidad de los píxeles se encuentra por un lado, el método planteado por Breeuwer et al en el año 2003. Este método, partiendo de imágenes en las que un especialista había marcado manualmente el epicardio y endocardio, discrimina entre el tejido dañado y sano realizando un procedimiento de umbralización automático simple. La umbralización usa el histograma de los píxeles del miocardio para determinar un umbral  $T$ . En primer lugar, se fija la intensidad en la que se encuentra la cima del histograma correspondiente al miocardio normal como la intensidad media ( $\mu$ ) del tejido sano. Seguido, se estima la desviación estándar ( $\sigma$ ) de las áreas sanas y finalmente define el umbral  $T$  de la siguiente manera:  $T = \mu + c \sigma$ , siendo  $c$  un valor ajustable por el usuario entre 1 y 6 [5].

Por otro lado, Amado et al. en el año 2004 propusieron otro enfoque de modelo fijo denominado Full-Width at Half-Maximum (FWHM), donde se selecciona como umbral de intensidad la mitad de la intensidad máxima dentro de una región seleccionada por el especialista [6]. Ambos métodos requieren de la intervención de un especialista, haciendo que los resultados puedan variar dependiendo de quién los ejecute. Con el objetivo de disminuir esta variación se han desarrollado otros modelos que dependen menos del usuario.

Un año más tarde Positano et al. (2005) desarrollaron un método para la definición de la extensión de la necrosis miocárdica en la que se extraían las regiones dañadas mediante el

método fuzzy c-mean (FCM) clustering, sin la necesidad de aplicar un umbral explícito [7]. Los FCM son un tipo de algoritmos de agrupamiento en los que cada elemento tiene un grado de pertenencia difuso a un grupo u otro. En general, el procedimiento de este método consiste en minimizar iterativamente una función objetivo hasta obtener una partición difusa óptima.

En 2008 Hennemuth [8], con el objetivo de visualizar la localización del tejido infartado, la forma y la transmuralidad de éste, presentó un método diferente. Éste está basado en algunas propiedades fundamentales del procesamiento de imágenes y de características de los infartos de miocardio. En primer lugar se define el umbral de intersección mediante un algoritmo de maximización de expectativas teniendo en cuenta el histograma y las distribuciones estadísticas que modelan éste. Una vez fijado el umbral, la segmentación se realiza con la técnica de watershed, técnica clásica que considera el gradiente de intensidad de la imagen como un área topográfica. En esta técnica, estructuras como las cicatrices pueden entenderse como un gradiente de intensidad alto en los bordes y menor en el centro, pudiendo ser posible segmentar el área que se encuentra entre los bordes.

Con la intención de mejorar los resultados obtenidos hasta la fecha, otros grupos de investigación presentaron modelos teniendo en cuenta más características además de la intensidad de los píxeles. Entre ellos el grupo de O'Donnell [9] planteó un modelo de detección en el que el miocardio se divide manualmente en 72 sectores circulares y se calcula de cada sector la intensidad media, el espesor de la sección interior y de la exterior, la homogeneidad y la transmuralidad. Una vez obtenidas las características y habiendo caracterizado un especialista a qué grupo (tejido sano / dañado) pertenece cada segmento, se entrena un clasificador *Support Vector Machine* (SVM) con todas las imágenes excepto con una, que se emplea para verificar el correcto funcionamiento del clasificador.

En el año 2013, Kotu et al. [10] presentó una técnica para describir mediante un mapa de probabilidades la naturaleza heterogénea del miocardio infartado. Para ello, se empleó una variedad de características más amplia que en las propuestas anteriores, contando con características de textura e intensidad. En los resultados de este estudio el tejido infartado y el tejido sano se representan por probabilidades altas y bajas, respectivamente, mientras que los valores intermedios indican áreas en las que se encuentran mezclados los tejidos cicatrizados y los sanos. El mapa de probabilidad se calculó empleando una función de probabilidad basada en la regla de Bayes.

Morisi et al. [11] publicó un artículo en el año 2014 en el que comparaba los resultados de clasificación obtenidos con diferentes técnicas de aprendizaje automático (SVM, redes neuronales, clasificador bayesiano y el método *k-nearest neighbors*). Todos los sistemas comparados tenían en cuenta diversas características, incluyendo entre ellas la forma, tamaño, brillo, contraste, etc. Su procedimiento comenzaba con la extracción inicial de todos los objetos similares a las cicatrices que se encontraban en el miocardio. Seguidamente, de cada uno de los objetos segmentados se calculaban 22 características (12 relacionadas con la forma y la geometría de los objetos y 10 basados en los niveles de gris). Una vez obtenidas las propiedades de cada objeto, los diferentes clasificadores se entrenaron con parte de ellas, empleando el resto en un testeo posterior.

Finalmente, Karim et. al. (2016) [2] ha presentado una review que recoge algunos de los algoritmos para la segmentación del tejido infartado generados en los últimos años y remarca lo difícil que es comparar los nuevos algoritmos con los ya existentes. Por ello, propone un marco de evaluación para los futuros algoritmos. Éste consiste en emplear los clasificadores desarrollados con una base de datos creada por ellos, la cual está compuesta por dos conjuntos de datos (humano y porcino), en los que los tejidos infartados están delimitados minuciosamente por 3 especialistas distintos que emplean el mismo protocolo. Una vez empleado el clasificador, los resultados tienen que ser comparados con el estándar de referencia creado por los especialistas mediante medidas de superposición y volumétricas. Con el objetivo de demostrar el buen funcionamiento de dicho marco, se evalúan cinco algoritmos siguiendo el mismo procedimiento.

## Objetivos

El objetivo principal de este trabajo es crear un sistema semiautomático que sea capaz de detectar tejido infartado en el miocardio analizando imágenes obtenidas mediante la técnica RTG por RMC y así reducir la carga de trabajo que supone en la actualidad para los especialistas la detección de dichos tejidos.

Con la intención de cumplir ese propósito, se ha tenido en cuenta el marco teórico revisado y se han llegado a las siguientes conclusiones:

En primer lugar, se ha entendido que además de la intensidad, existen más características que influyen en la caracterización de las regiones infartadas. Por lo tanto, el método propuesto también tiene en cuenta medidas relacionadas con la textura de la imagen.

En segundo lugar, visto que en el estudio mencionado en el apartado anterior de Morisi [11] se comprobó que la mejor solución obtenida, teniendo en cuenta la sensibilidad, la estabilidad y la dependencia del tamaño de muestras utilizado para entrenar, era el método SVM, se ha empleado dicho clasificador

Finalmente, con el propósito de conseguir una mayor exactitud se ha planteado trabajar a nivel de pixel en vez de realizarlo a nivel de regiones. Evitando así tener que hacer una segmentación previa de objetos que puedan ser regiones dañadas.

## Estructura del documento

**Capítulo 2:** Introducción de la base de datos empleada

**Capítulo 3:** Explicación del procedimiento realizado para la obtención de los resultados. Está dividido en diferentes apartados: El pre-procesado realizado, el cálculo de las características de cada pixel perteneciente al miocardio, el proceso de reducción de dimensionalidad llevado a cabo y la descripción de la creación y la implementación del clasificador

**Capítulo 4:** Exposición de los resultados preliminares obtenidos con el modelo generado evaluado mediante la validación cruzada *leave-one-out*.

**Capítulo 5:** Presentación del proceso de optimización llevado a cabo y los resultados finales obtenidos.

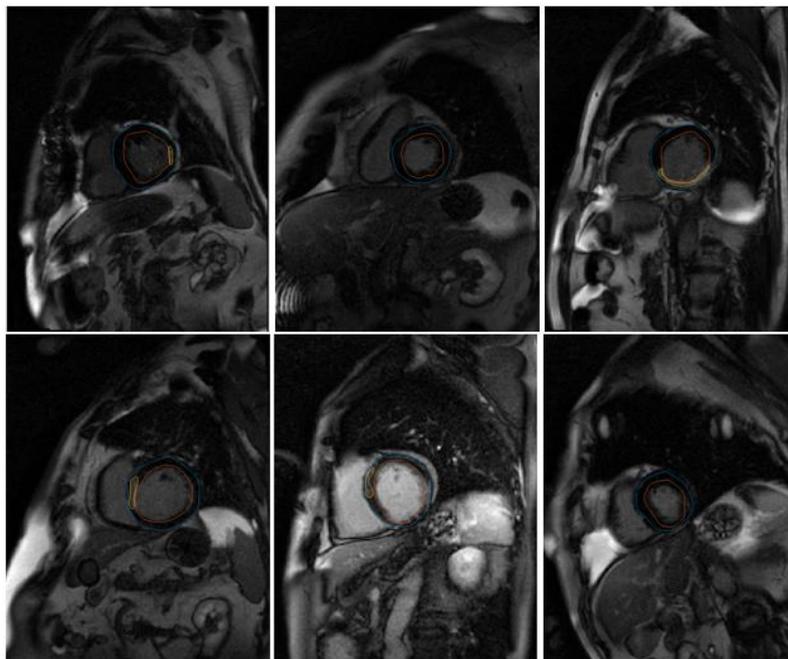
**Capítulo 6:** Agrupación de las conclusiones del trabajo realizado y las líneas futuras de investigación.

## Capítulo 2

# Base de datos

---

La base de datos empleada para la creación del clasificador ha sido cedida por el Hospital San Pedro, Rioja Salud de Logroño. Está compuesta de 20 estudios obtenidos entre los años 2013 y 2016. Los estudios pertenecen a diferentes pacientes que han sufrido al menos un infarto agudo de miocardio y cada uno contiene entre 7 y 11 cortes, dando un total de 164 imágenes. En cada una de las imágenes, un especialista ha definido manualmente las zonas de interés, el miocardio y las cicatrices en su caso. A continuación, (Fig 2) se muestran parte de las imágenes empleadas pertenecientes a diferentes estudios.



**Fig 2.** Imágenes de RMC pertenecientes a la base de datos empleada.

Las imágenes expuestas muestran como han sido marcados los bordes del miocardio que rodea el ventrículo izquierdo y el tejido infartado en caso de que haya. La línea amarilla define el área que cubre el tejido infartado y la línea roja y la azul señalan el borde interior y exterior del miocardio, respectivamente. El borde interior marcado es denominado endocardio (membrana que recubre el interior de las cavidades del corazón) mientras que el exterior puede ser epicardio (membrana epitelial que rodea todo el corazón) o el endocardio del ventrículo derecho, cuando se trata de la porción de miocardio que se encuentra entre los dos ventrículos. Con el objetivo de simplificar, en el trabajo se ha hecho referencia siempre al último borde presentado (línea azul) como epicardio.

El equipo de resonancia magnética empleado para la obtención de las imágenes ha sido el Siemens Avanto de 1.5T. Y la resolución de dichas imágenes es variable, siendo la mayoría de las imágenes de 256x200 y habiendo algún estudio compuesto por imágenes de resolución 216 x 256 y 144 x 192.

## Capítulo 3:

# Sistema propuesto

---

En este capítulo se va a introducir el sistema propuesto y los pasos que se han llevado a cabo para la creación del detector de tejido infartado.

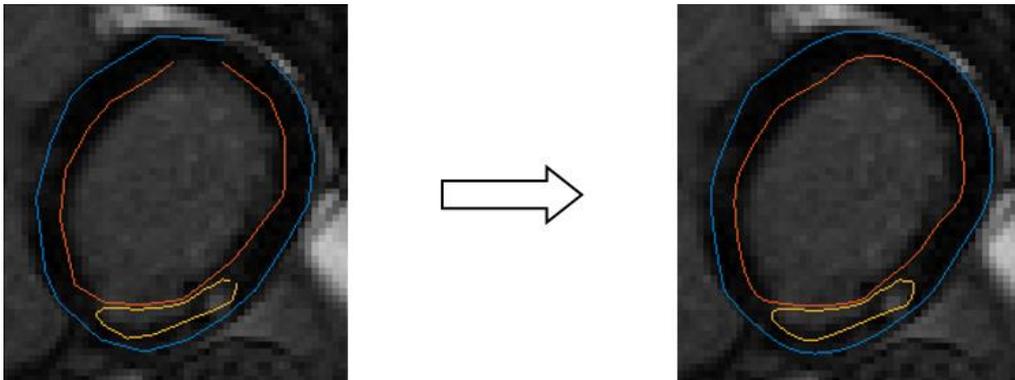
Primero, se ha realizado un leve pre-procesado extrayendo de las marcas realizadas por el especialista (miocardio y escara) todos aquellos píxeles pertenecientes al miocardio de cada una de las imágenes. A continuación, se ha calculado un conjunto de características de intensidad y textura que caracterizan a los píxeles del miocardio. Tras obtener estos parámetros se ha procedido a la reducción de la dimensionalidad de la matriz de características para finalmente, crear el clasificador, entrenándolo con parte de los píxeles del miocardio y testeándolo con el resto.

### Método de pre-procesado

Con el objetivo de reducir los tiempos de ejecución, se ha realizado un pre-procesado para extraer los píxeles del miocardio y calcular únicamente las características de dichos píxeles y no de los píxeles de toda la imagen. En futuras mejoras este paso deberá contar también con una previa segmentación automática del miocardio, sin que sea necesaria la ayuda del especialista, para así reducir más la carga que le supone a éste el análisis de los cortes.

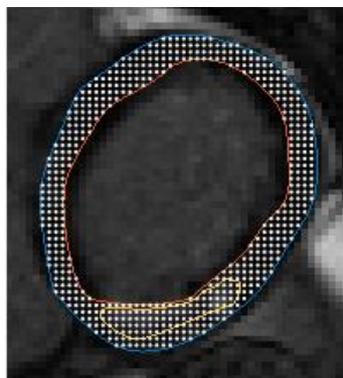
En primer lugar, se han interpolado las marcas que generaban los bordes del miocardio y del tejido infartado, pasando de los 19 puntos marcados por el especialista a 500. De esta manera,

como se puede apreciar en la siguiente imagen (Fig 3), se ha conseguido que los bordes sean continuos y más suaves.



**Fig 3.** Ejemplo de la interpolación de las marcas proporcionadas por el especialista.

Una vez interpolados los bordes, partiendo de las marcas generadas, se han seleccionado los píxeles que se encuentran en el interior de estos (Fig 4). Guardando a su vez los índices de los píxeles seleccionados en un vector que se empleará más adelante, asegurándose así de conocer siempre la localización de todos los píxeles.



**Fig 4.** Ejemplo de la selección de los píxeles pertenecientes al miocardio definido por el especialista.

Mediante el mismo procedimiento se han seleccionado los píxeles de tejido infartado; de esta manera se ha generado el vector de etiquetas que denomina con un 0 los píxeles de tejido sano y con un 1 los de tejido infartado.

## Extracción de características

A continuación van a ser explicados los diferentes operadores que han sido calculados para caracterizar los píxeles pertenecientes al miocardio de cada imagen. Entre los operadores se encuentran características estadísticas relacionadas con la intensidad o el nivel del gris de cada pixel, características de textura e incluso características geométricas. En este contexto, la textura se puede entender como la reproducción de un patrón espacial básico con una estructura periódica o aleatoria.

En el cálculo de todas las características en las que ha sido necesario se ha tenido en cuenta un vecindario de 3x3 píxeles.

### *Nivel de gris*

En primer lugar, se ha tenido en cuenta el nivel de gris del propio pixel normalizado con el nivel medio del miocardio. Como ya se ha comprobado en el estado del arte, esta característica resulta de gran interés ya que con la técnica de RTG por RMN los píxeles pertenecientes a los tejidos dañados son notoriamente de mayor intensidad que el resto, convirtiéndose así en una de las características más influyentes de la clasificación.

### *Nivel medio de gris*

El segundo parámetro calculado ha sido la media aritmética. Para el cálculo de esta se han tenido en cuenta los 8 píxeles vecinos más el propio pixel. La media aritmética por lo tanto se ha calculado como define la siguiente ecuación (1):

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{1}{n} \sum_{i=1}^n x_i \quad 1$$

Donde  $x_1, x_2, x_3, \dots, x_9$  son los niveles de gris de los 9 píxeles mencionados y  $n$  la cantidad de píxeles empleados, en este caso 9. Este parámetro aportará información del vecindario que rodea a cada pixel analizado.

### *Desviación estándar*

La desviación estándar también ha sido calculada para el vecindario de cada pixel. Esta medida da información sobre la dispersión y de lo alejados que están los valores del vecindario del nivel de gris medio del mismo.

Este parámetro es calculado como aparece en la ecuación 2 mediante la raíz cuadrada de la desviación cuadrática que tienen los niveles de gris de los píxeles respecto a la media de estos.

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \quad 2$$

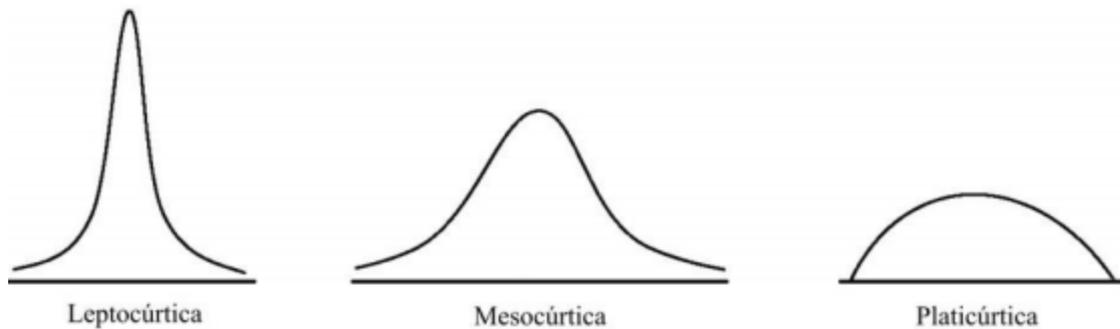
Siendo una vez más  $n$  igual a 9 y  $x_i$  los niveles de gris de los 9 píxeles que conforman el vecindario.

### ***Coeficiente de curtosis***

El coeficiente de curtosis mide el grado de apuntamiento que tienen los datos respecto a la distribución normal de referencia y se define de la siguiente manera (3):

$$Curt = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{S^4} - 3 \quad 3$$

Si el valor del parámetro es igual a 0, la distribución será *mesocúrtica*, lo que corresponde a un grado de apuntamiento como el que tiene la curva de distribución normal. En cambio, si el valor de la curtosis es negativo la curva será *platocúrtica* y si la curtosis es positiva, la curva será *leptocúrtica* (Fig 5).



**Fig 5.** Formas y clasificación de las curvas de distribución, según la *curtosis*.

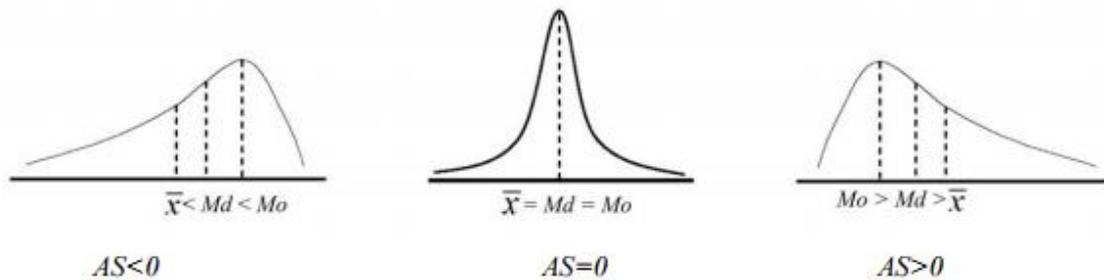
Empleando este parámetro se obtiene información sobre el grado de concentración que tienen los niveles de gris pertenecientes al vecindario en torno a la media aritmética de este.

### ***Coeficiente de asimetría de Fisher***

El coeficiente de Fisher es una medida de asimetría de la disposición de los datos con respecto a la media aritmética. Se define mediante la siguiente expresión matemática (4):

$$AS = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{S^3} \quad 4$$

Si el parámetro es positivo, los valores del vecindario se distribuyen más hacia la derecha de la media de este, por lo tanto habrá mayor dispersión de los datos superiores a la media, siendo asimetría derecha o positiva. Mientras que si el valor es negativo, los valores se distribuyen más hacia la izquierda, consiguiendo una asimetría izquierda o negativa. En caso de que se trata de una distribución simétrica o normal, el valor del coeficiente será igual a cero (Fig 6).



**Fig 6.** Formas de la curva de distribución según la simetría. Md es la media y Mo representa la moda, el valor que presenta una mayor frecuencia.

### **Matriz de Co-ocurrencia de Nivel de Gris (GLCM)**

La Matriz de co-ocurrencia de nivel de gris (GLCM) [12] es una medida de textura que da información sobre la disposición espacial de las intensidades del vecindario escogido. Este parámetro describe la frecuencia de un nivel de gris que aparece en una relación espacial específica con otro valor de gris, dentro del área de una ventana determinada.

La definición más simple de la probabilidad de la matriz es: “el número de veces que un evento ocurre, dividido por el número total de posibles eventos” [13] y la ecuación para su cálculo es la siguiente (5):

$$P_d(i, j) = \frac{C_d(i, j)}{\sum_i \sum_j C_d(i, j)} \quad 5$$

Siendo  $i$  y  $j$  el número de filas y columnas respectivamente y  $C_d(i, j)$  el número de veces que aparece el par de píxeles  $(i, j)$  cuando los píxeles se encuentran separados por la distancia  $(1;0)$  en diferentes orientaciones. En la Fig 7 se puede apreciar un ejemplo del cálculo de la matriz realizado teniendo en cuenta la orientación vertical.

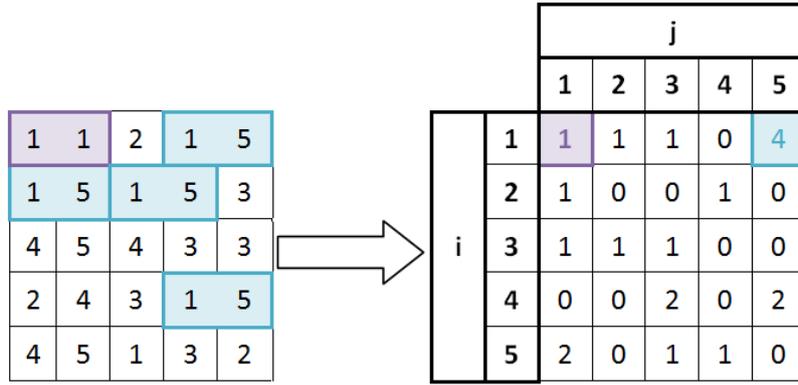


Fig 7. Ejemplo del cálculo de una GLCM en dirección vertical.

A pesar de que esta característica capture las propiedades de la textura, no resulta directamente útil para un análisis posterior en el que se comparen dos texturas. Por lo tanto, a partir de la GLCM normalizada se han calculado otras características numéricas como son la energía, el contraste, la homogeneidad, la correlación y la entropía, definidas mediante las siguientes ecuaciones respectivamente (6, 7, 8, 9 y 10):

$$GLCM \text{ Energía} = \sum_{i,j=0}^{N-1} P_{i,j}^2 \quad 6$$

$$GLCM \text{ Contraste} = \sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2 \quad 7$$

$$GLCM \text{ Homogeneidad} = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \quad 8$$

$$GLCM \text{ Correlación} = \sum_{i,j=0}^{N-1} P_{i,j} \left[ \frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right] \quad 9$$

$$GLCM \text{ Entropía} = \sum_{i,j}^{N-1} -P_{i,j} \ln(P_{i,j}) \quad 10$$

Siendo  $P_{i,j}$  la probabilidad de co-ocurrencia de los valores de gris  $i$  y  $j$ , para una distancia dada y en la correlación  $\mu_i$  y  $\mu_j$  las medias y  $\sigma_i$  y  $\sigma_j$  las varianzas del nivel de gris de  $i$  y  $j$ , respectivamente [A9].

## Local Binary Pattern (LBP)

El *Local Binary Pattern* [12] es un tipo de operador de textura eficiente utilizado especialmente para enfatizar la textura de la imagen en pequeñas escalas y destacar similitudes entre las imágenes. Teóricamente, considerando el nivel de gris de un píxel arbitrario  $g_c$  de una imagen, el patrón binario local reemplaza cada valor de  $g_c$  por un código binario de 8 bits. Siendo las variables de entrada el nivel de gris del píxel central  $g_c$ , el radio elegido  $R$  y el número de píxeles vecinos  $P$ . Generalmente, con el fin de modelar la textura de la imagen a pequeña escala, se escoge una ventana de radio 1 ( $R = 1 = 3 \times 3$ ) centrada en  $g_c$  con sus  $P=8$  píxeles vecinos.

El método consiste en calcular la diferencia entre el píxel central  $g_c$  y sus  $P$  píxeles vecinos ( $g_p, p = 0, \dots, P - 1$ ). A continuación se aplica una función de umbral  $s(x)$ . Si el valor de la diferencia es menor que 0, se asignará 0. De lo contrario, 1. Finalmente, la imagen LBP correspondiente se obtiene traduciendo el código binario en su valor decimal. Se sigue una dirección en el sentido de las agujas del reloj, comenzando con el píxel en el lado derecho de  $g_c$ . Esta característica puede ser expresada matemáticamente por las siguientes ecuaciones (11 y 12):

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad 11$$

$$s(x) = \begin{cases} 1 & \text{si } (g_p - g_c) \geq 0 \\ 0 & \text{si } (g_p - g_c) < 0 \end{cases} \quad 12$$

A continuación (Fig 8) se muestra un ejemplo práctico del cálculo de este operador de textura.

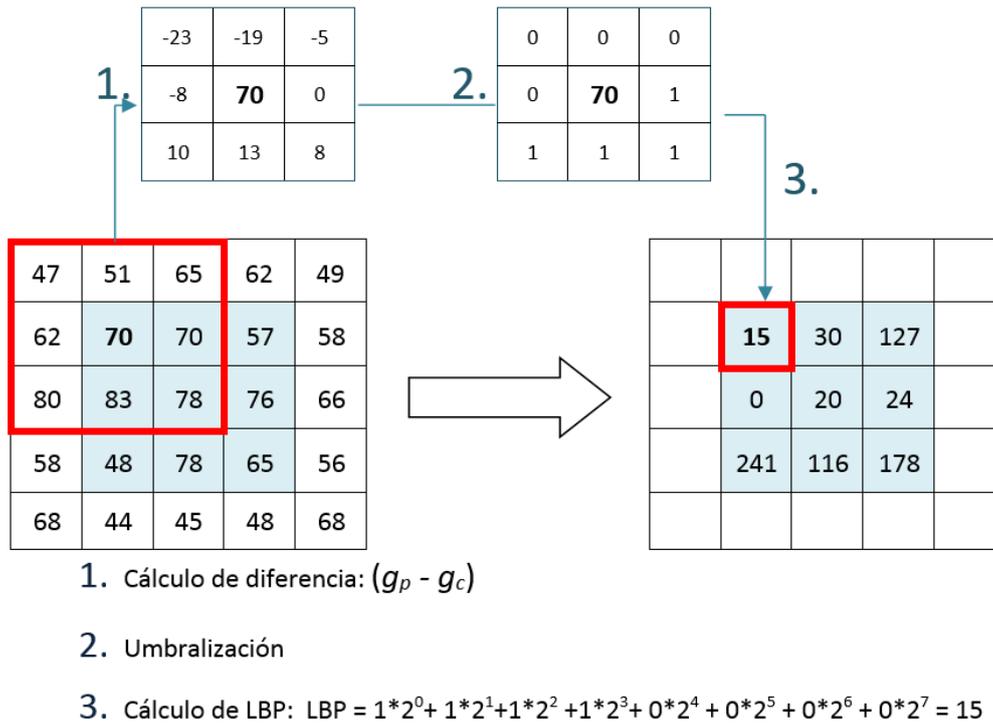


Fig 8. Ejemplo del proceso llevado a cabo para el cálculo de LBP.

En este caso, además de tener en cuenta para cada pixel el valor de LBP que le pertenece también se han guardado los valores de LBP correspondientes a sus 8 vecinos; en el ejemplo anterior, para el pixel central se guardarían los píxeles marcados de azul.

Una vez obtenidos los 9 valores se ha repetido el proceso empleando un R de 2 y guardando finalmente los valores de LBP de los 25 píxeles que componen el nuevo vecindario tenido en cuenta esta vez.

### Laws Texture Energy (LTE)

El LTE [12] es otro método que proporciona características de textura. En este caso, se emplean máscaras locales para detectar diferentes tipos de textura de los que se estimará la energía. Para el cálculo se han empleado un conjunto de tres máscaras de una dimensión convolucionadas entre sí. Las máscaras unidimensionales, de tamaño 3, se denominan L3, E3 y S3 (12), y describen el nivel, el borde y la característica del punto respectivamente.

$$\begin{aligned}
 L3 &= [1,2,1] \\
 E3 &= [-1,0,1] \\
 S3 &= [-1,2,-1]
 \end{aligned}
 \tag{12}$$

Como se puede apreciar en la siguiente imagen (Fig 9) mediante la convolución de estas máscaras entre sí, se obtienen nueve nuevas máscaras de dos dimensiones (3x3). Se han realizado todas las combinaciones posibles: L3L3 (mask1), S3S3 (mask2), E3E3 (mask3), S3L3 (mask4), L3S3 (mask5), L3E3 (mask6), E3L3 (mask7) E3S3 (mask8) y S3E3 (mask9).

Mask1: L3L3	Mask2: S3S3	Mask3: E3E3	Mask4: S3L3	Mask5: L3S3																																													
<table border="1"><tr><td>1</td><td>2</td><td>1</td></tr><tr><td>2</td><td>4</td><td>2</td></tr><tr><td>1</td><td>2</td><td>1</td></tr></table>	1	2	1	2	4	2	1	2	1	<table border="1"><tr><td>1</td><td>-2</td><td>1</td></tr><tr><td>-2</td><td>4</td><td>-2</td></tr><tr><td>1</td><td>-2</td><td>1</td></tr></table>	1	-2	1	-2	4	-2	1	-2	1	<table border="1"><tr><td>1</td><td>0</td><td>-1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>-1</td><td>0</td><td>1</td></tr></table>	1	0	-1	0	0	0	-1	0	1	<table border="1"><tr><td>-1</td><td>-2</td><td>-1</td></tr><tr><td>2</td><td>4</td><td>2</td></tr><tr><td>-1</td><td>-2</td><td>-1</td></tr></table>	-1	-2	-1	2	4	2	-1	-2	-1	<table border="1"><tr><td>-1</td><td>2</td><td>-1</td></tr><tr><td>-2</td><td>4</td><td>-2</td></tr><tr><td>-1</td><td>2</td><td>-1</td></tr></table>	-1	2	-1	-2	4	-2	-1	2	-1
1	2	1																																															
2	4	2																																															
1	2	1																																															
1	-2	1																																															
-2	4	-2																																															
1	-2	1																																															
1	0	-1																																															
0	0	0																																															
-1	0	1																																															
-1	-2	-1																																															
2	4	2																																															
-1	-2	-1																																															
-1	2	-1																																															
-2	4	-2																																															
-1	2	-1																																															
Mask6: L3E3	Mask7: E3L3	Mask8: E3S3	Mask9: S3E3																																														
<table border="1"><tr><td>-1</td><td>0</td><td>1</td></tr><tr><td>-2</td><td>0</td><td>2</td></tr><tr><td>-1</td><td>0</td><td>1</td></tr></table>	-1	0	1	-2	0	2	-1	0	1	<table border="1"><tr><td>-1</td><td>-2</td><td>-1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>1</td><td>2</td><td>1</td></tr></table>	-1	-2	-1	0	0	0	1	2	1	<table border="1"><tr><td>1</td><td>-2</td><td>1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>-1</td><td>2</td><td>-1</td></tr></table>	1	-2	1	0	0	0	-1	2	-1	<table border="1"><tr><td>1</td><td>0</td><td>-1</td></tr><tr><td>-2</td><td>0</td><td>2</td></tr><tr><td>1</td><td>0</td><td>-1</td></tr></table>	1	0	-1	-2	0	2	1	0	-1										
-1	0	1																																															
-2	0	2																																															
-1	0	1																																															
-1	-2	-1																																															
0	0	0																																															
1	2	1																																															
1	-2	1																																															
0	0	0																																															
-1	2	-1																																															
1	0	-1																																															
-2	0	2																																															
1	0	-1																																															

Fig 9. Mascaras de LTE empleadas.

Debido a que todas excepto la primera máscara (L3L3) tienen una media de cero, la primera máscara (L3L3) se ha utilizado para normalizar el contraste de todas las demás imágenes de textura, mientras que el resto se han empleado por separado para el cálculo de las diferentes imágenes de textura.

Como se ha comentado anteriormente, la textura se considera como un conjunto de píxeles primitivos que mantienen una relación regular o repetida. Por lo tanto, para la obtención de la imagen de textura (TI, ecuación 13) que detecte los diferentes tipos de textura, es necesario realizar una convolución de la imagen original con todas las máscaras locales definidas por separado. Obteniendo como resultado de este proceso una imagen completa que resalta las texturas definidas por la máscara con la que ha sido convolucionada. A continuación se explica matemáticamente como se realizaría el proceso empleando como ejemplo la máscara2 (S3S3):

$$TI_{S3S3} = I * S3S3 \quad 13$$

$$TI_{S3S3 \text{ Normaliz.}} = \frac{TI_{S3S3}}{TI_{L3L3}} \quad 14$$

Siendo \* el operador de convolución. Las imágenes de TI normalizadas resultantes se pasan a través de un filtro medio de valores absolutos, obteniéndose el mapa de energía de textura (TE) de la siguiente manera (15):

$$TE(i, j) = \sum_{-3}^3 \sum_{-3}^3 |TI_{i+u, j+v}| \quad 15$$

Siendo  $u, v$  las dimensiones de la imagen.

### *Gabor wavelet (GW)*

A través de la Transformación Gabor Wavelet (GW), los filtros de Gabor proporcionan tanto la frecuencia (estructura estacionaria y periódica) como la localización espacial. Esto se puede lograr mediante la convolución de la imagen con filtros de Gabor de diferentes escalas y orientaciones, como los que se muestran en la Fig 10. Este tipo de parámetros permiten el análisis de la variación espacial de una manera similar al sistema visual humano, siendo una herramienta útil para la discriminación de textura.

En el dominio espacial, un filtro Gabor de dos dimensiones es un kernel Gaussiano o una envolvente modulada por una onda sinusoidal a lo largo del eje  $x$ . Este filtro está compuesto por una componente real y otra imaginaria. El filtro Gabor,  $gf$ , puede ser definido de la siguiente manera (16):

$$gf_{(l,k)}(m, n) = \frac{f^2}{\pi\gamma\eta} \exp\left(-x'^2 \frac{f^2}{\gamma^2} + y'^2 \frac{f^2}{\eta^2}\right) \exp(j2\pi f x') \quad 16$$

Donde  $f$  es la frecuencia del factor sinusoidal,  $\gamma$  la nitidez a lo largo del eje mayor  $X$ , y  $\eta$  la nitidez a lo largo del eje menor  $Y$ . Por otro lado,  $l$  y  $k$  son los números enteros que identifican la escala y el factor de orientación respectivamente,  $m$  y  $n$  son las dimensiones de la imagen y finalmente  $x'$  e  $y'$  se definen del siguiente modo (17 y 18):

$$x' = m \cos(\theta) + n \sin(\theta) \quad 17$$

$$y' = -m \sin(\theta) + n \cos(\theta) \quad 18$$

Siendo  $\theta$  la orientación de la normal a las franjas paralelas de la función de Gabor.

Normalmente, se emplea un conjunto de filtros Gabor Gaussianos de diferentes tamaños modulados por ondas planas sinusoidales de diferentes orientaciones. El tamaño del núcleo es proporcional al ancho del Gaussiano. La representación del GW de una imagen es la convolución de la imagen con una serie de GW. En este trabajo se ha empleado un conjunto de 24 GW

complejas como las que se pueden ver en la Fig 10, de 4 escalas y 6 orientaciones (0°, 30°, 60°, 90°, 123°, 150°).

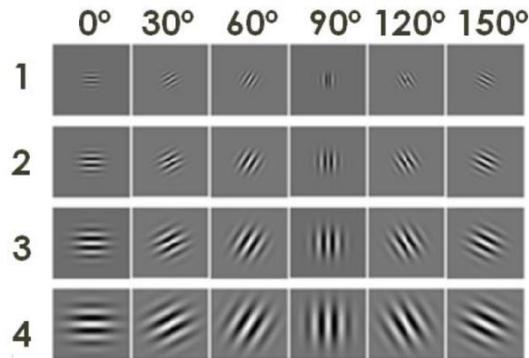


Fig 10. Filtros de GW empleados.

Matemáticamente los resultados se pueden obtener de la siguiente manera (17):

$$gI_{(l,k)} = I * gf_{(l,k)} \quad 17$$

En este caso, en el vector de características para cada pixel se ha guardado el propio valor de Gabor calculado teniendo en cuenta un vecindario de 3x3 píxeles y la media y la desviación estándar de todos estos valores en el mismo vecindario. Calculándose la media y la desviación de la siguiente manera (18 y 19):

$$\mu_{l,k}(m, n) = \frac{1}{M \cdot N} \sum_{m=1}^M \sum_{n=1}^N |gI_{l,k}(m, n)| \quad 18$$

$$\sigma_{l,k}(m, n) = \sqrt{\frac{1}{M \cdot N} \sum_{m=1}^M \sum_{n=1}^N (|gI_{l,k}(m, n)| - \mu_{l,k}(m, n))^2} \quad 19$$

## Distancias

Por último se han calculado también las distancias que hay; por un lado, entre el pixel y el epicardio y por el otro, del pixel al endocardio. Ambas distancias han sido normalizadas respecto al grosor que tiene el miocardio en ese tramo. Se trata de una característica interesante ya que tras una oclusión, la evolución del tejido infartado es desde la cavidad (tejido cercano al endocardio) hacia la transmuralidad (todo el espesor de la pared del miocardio). Es decir, en las primeras fases el tejido infartado afecta solamente las áreas más cercanas al endocardio y conforme pasa el tiempo el tejido afectado va aumentando hasta dañar todo el espesor del

miocardio en la zona dañada. En la siguiente imagen se representa el proceso de manera visual (Fig 11).

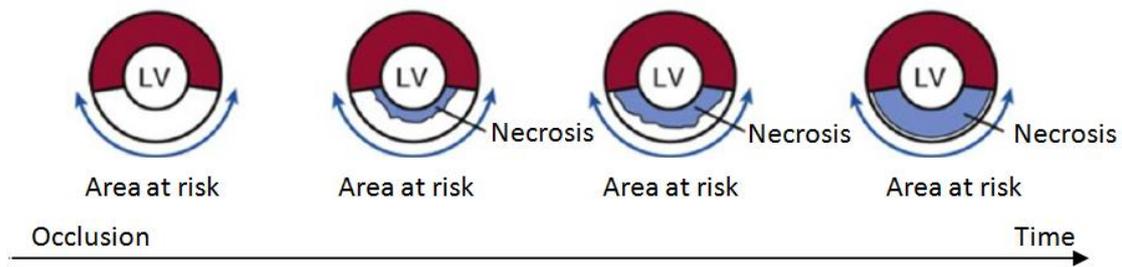


Fig 11. Evolución del tejido infartado o la necrosis con el tiempo.

## Selección de características

Este apartado se centra en la reducción de dimensionalidad de la matriz de características. Es decir, en el proceso de reducir el número de variables o características empleadas en el entrenamiento del clasificador. De esta manera, se ha evitado trabajar con características redundantes que podrían afectar en los clasificadores al lidiar con un problema de dimensionalidad. Esta reducción también puede ser beneficiosa en cuanto al tiempo de almacenamiento y computación.

### Variables

Tras calcular las características se han obtenido tres variables:

- **Matriz de características:** Consiste en una matriz de 82.166 observaciones correspondientes a todos los píxeles de los miocardios de las 164 imágenes y 134 características para cada pixel, siendo estas las características:
  - 1 Nivel de gris normalizado
  - 4 estadísticos: Media aritmética, desviación estándar, coeficiente de curtosis y coeficiente de asimetría (dim. 3x3)
  - 5 GLCM: Energía, contraste, homogeneidad, correlación y entropía (dim. 3x3)
  - 34 LBP: propio y de los vecinos (dim. 3x3 y 5x5)
  - 16 LTE
  - 72 GW: (4 escalas y 6 orientaciones): 24 GW, 24 GW media y 24 GW desviación estándar (dim. 3x3)
  - 2 distancias: Pixel – epicardio y pixel – endocardio.

- **Label:** define a qué grupo pertenecen los 82.166 píxeles. 0 corresponde al grupo de tejido sano mientras que 1 al grupo de tejido infartado.
- **Índices:** define la posición en la que se encuentra cada pixel. Por un lado, se encuentra el índice que hace referencia a la localización del pixel en la imagen y por otro el que hace referencia a la imagen a la que pertenece dicho pixel.

### *Prueba de t-test*

El tipo más común de la prueba t-test es la prueba de t de Student. En este caso se ha empleado la prueba de t de Student para dos poblaciones. El objetivo de esta prueba es la de evaluar si existen diferencias significativas entre las medias de las poblaciones de interés, es decir entre los datos provenientes del tejido sano y las del tejido infartado. A la hora de realizar el análisis se tiene en cuenta las siguientes suposiciones:

- Cada una de las dos poblaciones comparadas sigue una distribución normal. Esta condición puede ser comprobada realizando un test de normalidad.
- Las dos poblaciones que van a ser comparadas deben tener la misma varianza. Sin embargo, si el tamaño de las poblaciones a comparar es el mismo, la presencia de varianzas diferentes no afecta en el test.

El teorema del límite central [14] establece que, dadas ciertas condiciones muy generales, la media aritmética de una muestra aleatoria de una población con varianza finita, tiene una distribución aproximada a la normal cuando el tamaño de la muestra es grande (superior a 30), independientemente de la forma de la distribución de la población. En este caso el tamaño de la muestra contiene un número superior a 30 observaciones, siendo 74.834 píxeles de tejido sano y 7.332 de tejido infartado. Por lo tanto, se cumplen las suposiciones exigidas y se puede continuar con la implementación de la prueba t de Student.

### *Implementación de t-test*

A la hora de implementar la prueba se ha entendido como hipótesis nula que las medias de ambas poblaciones son iguales, mientras que la hipótesis alternativa indica que las medias no son iguales. El nivel de significación empleado para esta prueba ha sido del 1%.

Una vez definidas las hipótesis y el nivel de significación se ha llevado a cabo la prueba, definiendo como características significativas aquellas que tienen como resultado un p-valor menor al nivel de significación, ya que con las características que cumplan esta condición habrá evidencias suficientes para rechazar la hipótesis nula, resultando ser las medias de las dos poblaciones (tejido sano y tejido infartado) diferentes.

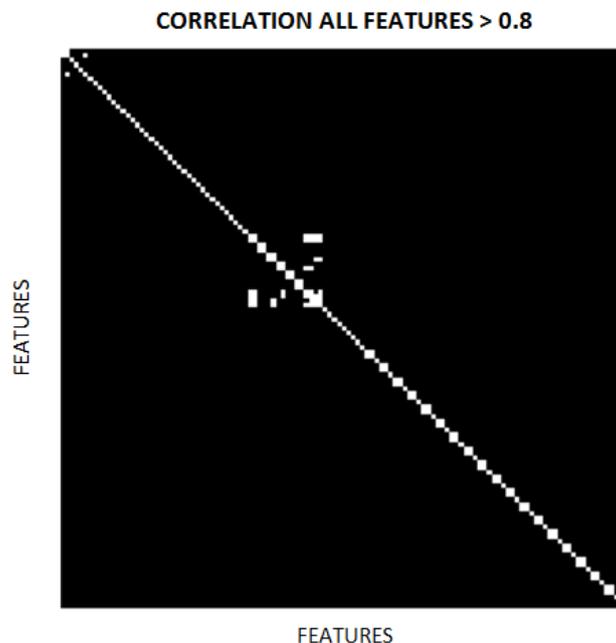
Realizado este proceso, se ha conseguido reducir la matriz de la que se partía a otra que cuenta con 121 características, todas ellas significativas a la hora de comparar las dos poblaciones.

### *Análisis de correlación*

Tras eliminar las características que no son significativas para diferenciar entre los dos tejidos, se ha realizado un análisis de correlación de los datos conservados. Este estudio permite conocer la existencia de características relacionadas entre sí, es decir anuncia la existencia de dependencias lineales entre dos variables.

Esta medida toma valores entre -1 y 1. Considerándose que no hay relación lineal entre las variables cuando el valor del coeficiente de correlación es igual a 0. Mientras que cuanto mayor sea el valor absoluto del coeficiente la relación lineal entre las dos variables será más fuerte.

Para el análisis se han comprobado todas las combinaciones posibles entre todas las características. Debido a la gran cantidad de variables con las que se está trabajando, en vez de mostrar la gráfica de matriz en la que se observarían todas las relaciones comentadas se ha extraído una imagen (Fig 12) en la que se muestran las que tienen una correlación absoluta mayor que 0.8 partiendo de las 121 características resultantes del análisis t de Student.



**Fig 12.** Representación de características correlacionadas con un coeficiente mayor que 0.8.

En la figura anterior se marcan con un punto blanco aquella pareja de características que se considera que están relacionadas linealmente entre sí, ya que tienen un coeficiente de

correlación mayor que 0.8. En la imagen se observa una diagonal que representa la relación de las características con ellas mismas.

De esta manera, se ha observado que existen entre las características varios casos de correlación, por ello se han analizado todas las correlaciones y se ha optado por eliminar aquellas características que resulten redundantes con el objetivo de reducir los casos de correlación. Mediante este procedimiento han sido reducidas las características a las siguientes 72:

- 1 Nivel de gris normalizado
- 4 estadísticos: Media aritmética, desviación estándar, coeficiente de curtosis y coeficiente de asimetría (dim. 3x3)
- 4 GLCM: Contraste, homogeneidad, correlación y entropía (dim. 3x3)
- 9 LBP
- 8 LTE
- 44 GW
- 2 distancias: Pixel – epicardio y pixel – endocardio.

### *Simplificación final de las características*

Finalmente, tras generar el clasificador y con el objetivo de reducir el peso de los datos y optimizar el funcionamiento de la herramienta, se ha vuelto a modificar el número de características, eliminando todas aquellas que no mejoraban el resultado. Para ello se han ido eliminando diferentes características, se ha estudiado su efecto en el resultado final y han sido descartadas aquellas que apenas tienen influencia en el resultado, reduciendo así el número de características a un conjunto de 48.

### *Estandarización*

La normalización es un método utilizado para estandarizar el rango de variables independientes. Dado que el rango de los valores empleados en este trabajo varía ampliamente, existe la necesidad de normalizar las características.

Durante la prueba de t-test no es necesario estandarizar las características, ya que cada característica se compara individualmente. Sin embargo, será necesario realizar la estandarización antes de insertar los datos en el clasificador.

Pueden diferenciarse dos métodos principales:

1. Escalado: consiste en escalar los datos al rango [0, 1]. Este procedimiento se consigue con la ecuación presentada a continuación (20), donde  $x$  es el valor original y  $x'$  el valor modificado.

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad 20$$

2. Normalización: consiste en hacer que los valores de cada característica tengan una media de cero y una varianza de uno. El método consiste en determinar la media ( $\mu(x)$ ) y la desviación estándar ( $\sigma(x)$ ) para cada característica. A continuación, como muestra la ecuación 21 se resta al valor original ( $x_i$ ) la media y se dividen los valores de cada característica por su desviación estándar.

$$x'_i = \frac{x_i - \mu(x)}{\sigma(x)} \quad 21$$

Ambos métodos han sido empleados a lo largo del trabajo.

## *Organización de las características*

Finalmente, de las 134 características originales calculadas se han reducido las dimensiones en diferentes procesos hasta llegar a obtener una suma de 48 características, siendo las siguientes:

- 1 Nivel de gris normalizado
- 4 estadísticos: Media aritmética, desviación estándar, coeficiente de curtosis y coeficiente de asimetría (dim. 3x3)
- 4 GLCM: Contraste, homogeneidad, correlación y entropía (dim. 3x3)
- 9 LBP: propio y de los vecinos (dim. 3x3)
- 8 LTE
- 20 GW: 18 GW (3 escalas y 6 orientaciones), 1 GW media y 1 GW desviación estándar, significativas en la escala menor.
- 2 distancias: Pixel – epicardio y pixel – endocardio.

Estas 48 características obtenidas para todos los píxeles han sido guardadas en una sola matriz, en la que cada columna corresponde a un pixel y cada fila a una característica. Con el objetivo de facilitar el control de las etiquetas y de la localización de los píxeles, a dicha matriz se le han añadido las matrices índices y label explicadas anteriormente.

De esta manera se ha trabajado con una matriz como la que se muestra en la siguiente imagen (Fig 13) de 51 x 82.166 dimensiones:

	Image 1				Image 2				...	Image 164			
	P1	P2	...	PN	P1	P2	...	PN	...	P1	P2	...	PN
Feature 1									...				
Feature 2									...				
Feature 3									...				
...	...	...	...	...	...	...	...	...	...	...	...	...	...
Feature 48													
LABEL	0	1	0	0	0	1	0	1	...	1	0	0	1
Index									...				
Image	1	1	1	1	2	2	2	2	...	164	164	164	164

Fig 13. Organización de las características.

## Clasificación de cicatrices

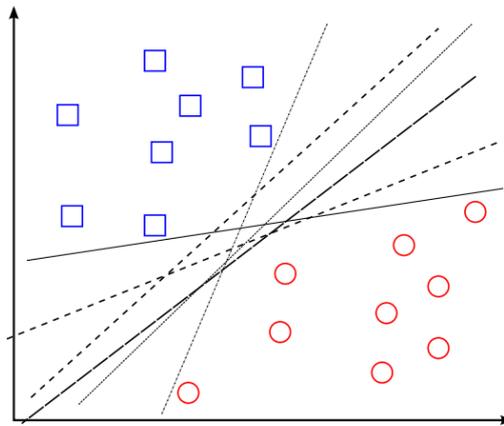
### *Support Vector Machine (SVM)*

El sistema de clasificación empleado ha sido una máquina de vector soporte, también llamado SVM del inglés Support Vector Machine. Se trata de un clasificador que mediante un grupo de entrenamiento previamente etiquetado y clasificado aprende el lugar en el que se encuentra la superficie de decisión entre dos clases.

La teoría de las SVM fue desarrollada en los años 90 por Vapnik y su equipo en los laboratorios AT&T (American Telephone & Telegraph Company). En un principio estas máquinas fueron pensadas para resolver problemas de clasificación binaria, pero en la actualidad se emplean para otros problemas también, como el agrupamiento, la regresión y la multclasificación.

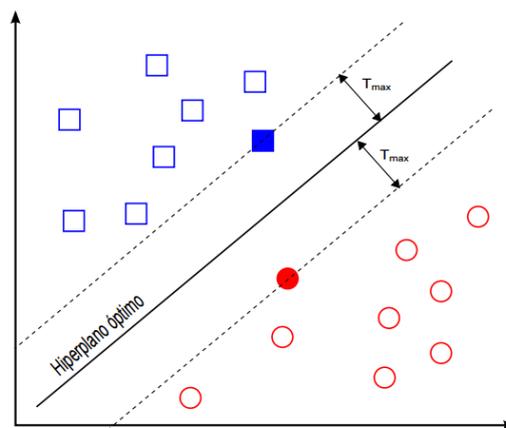
Las SVM se encuentran dentro de la categoría de clasificadores lineales, ya que generan separadores lineales o hiperplanos en espacios de características de alta dimensionalidad. Estos clasificadores en caso de ser necesario, son capaces de mapear los puntos de partida a un espacio de características de mayor dimensión y acabar encontrando el hiperplano que separa los dos grupos.

A continuación se muestra un ejemplo simple en el que se observa como la SVM calcula el hiperplano óptimo para un espacio linealmente separable. Para ello, en primer lugar la máquina conforma diferentes hiperplanos que separan los datos de entrada en los dos grupos que poseen una etiqueta propia (Fig 14).



**Fig 14.** Ejemplo de obtención de posibles hiperplanos de separación en un conjunto separable en dos clases.

Entre todos los planos de separación posibles, existe un solo hiperplano de separación óptimo. Este hiperplano será aquel que contenga el margen máximo. Siendo el margen ( $\tau$ ) la mínima distancia entre el hiperplano de separación y los datos más cercano de las dos clases (uno de cada clase). Por lo tanto, el hiperplano óptimo en el ejemplo anterior será el que aparece en la siguiente figura (Fig 15).



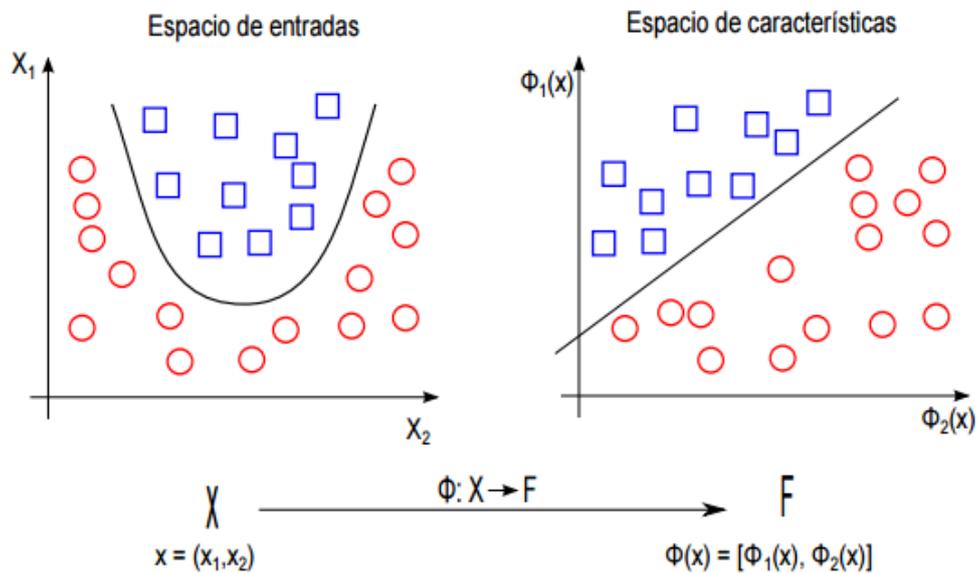
**Fig 15.** Hiperplano de separación óptimo logrado con el margen asociado máximo en el ejemplo anterior.

Los puntos que se tienen en cuenta a la hora de crear el margen máximo (puntos coloreados en la Fig 15) son denominados vectores de soporte.

Este caso sería solamente útil para trabajar con clases perfectamente separables o cuasi-separables linealmente. Cuando se trate de grupos no separables linealmente en el espacio de entrada pero si en otro espacio diferente, es necesario modificar el planteamiento comentado.

En estos casos, es preciso realizar una transformación previa de las variables de entrada mediante una función kernel.

Una función kernel o núcleo es un producto interno del espacio de características equivalente al espacio de entrada. Gráficamente se puede observar el proceso en la siguiente imagen (Fig 16) que ilustra un ejemplo de cómo la función kernel, partiendo del espacio de entrada en el que no se separan las clases linealmente, permite reubicar los datos a un espacio de características en el que es posible realizar la separación con el método de margen máximo aplicado en los casos lineales [15].



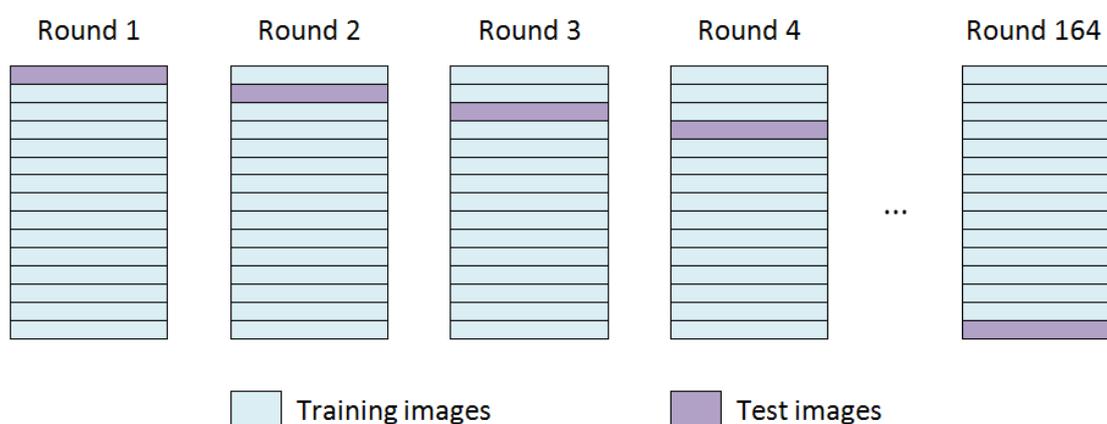
**Fig 16.** Ejemplo de la transformación y separación realizada por SVM a un grupo de entrada no separable linealmente.

Entre los kernels más comunes se encuentran los lineales, los polinómicos y los gaussianos.

### *Evaluación del clasificador*

Los pixeles del vector de características han sido usados para entrenar el clasificador. Con el fin de tener una indicación de lo bueno que es el clasificador cuando vaya a tratar nuevas predicciones, a la hora de entrenar no se han utilizado todos los datos. Los que han sido eliminados se han empleado posteriormente para probar el rendimiento del clasificador. Por lo tanto, los datos han sido divididos en dos conjuntos: entrenamiento y testeo. Los de entrenamiento han sido empleados para entrenar el clasificador, mientras que los de testeo (las muestras restantes) se han utilizado para evaluar su rendimiento. Este método de evaluación es denominado validación cruzada (cross-validation).

En este estudio se ha empleado la validación cruzada *leave-one-out*. Esta validación implica separar los datos de manera que en cada iteración se tenga una sola muestra en el grupo de testeo y todo el resto forme parte del grupo de entrenamiento. En este caso, en una prueba se emplearán como muestra única en el grupo de testeo todos los píxeles pertenecientes al miocardio de una imagen, y en una segunda los de un estudio completo. El procedimiento se repite 164 veces (número de imágenes con las que se cuenta) y cada vez, los píxeles de una imagen se emplearán como el conjunto de test y el resto se recogerán para formar un conjunto de entrenamiento (Fig 17). A continuación, es calculada la precisión de la validación de todas las imágenes para obtener una precisión de validación cruzada final.



**Fig 17.** Ejemplo de validación cruzada.

A la hora de probar cuantitativamente el rendimiento del modelo desarrollado se han tenido en cuenta diferentes parámetros. Dado que los tejidos que clasifique como sano o dañado pueden ser tanto verdaderos como falsos se han denominado de la siguiente manera las cuatro posibilidades que se pueden dar a la hora de la clasificación:

- TP, Verdadero Positivo: Píxeles que han sido clasificados correctamente como tejido dañado.
- TN, Verdadero Negativo: Píxeles clasificados correctamente como tejido sano.
- FP, Falso Positivo: Píxeles clasificados erróneamente como tejido dañado.
- FN, Falso Negativo: Píxeles clasificados erróneamente como tejido sano.

Los parámetros tenidos en cuenta para evaluar cuantitativamente el rendimiento del modelo son los siguientes:

- Accuracy ( $Acc$ ): Medida que da información de cómo de bien identifica el clasificador los datos en la clase correspondiente. Es decir, la proporción de resultados verdaderos (tanto de tejido sano como infartado) entre el número total de casos (ecuación 22). Esta medida se presenta en porcentaje.

$$ACC = \frac{\sum TP + TN}{\sum Total\ Population\ (TP + TN + FP + FN)} \quad 22$$

- Sensibilidad ( $Sn$ ): Proporción de píxeles dañados que se identifican correctamente (ecuación 23). Se presenta en porcentaje.

$$Sn = \frac{\sum TP}{\sum Condition\ Positive\ (TP + FN)} \quad 23$$

- Especificidad ( $Sp$ ): Proporción de píxeles sanos correctamente identificados (ecuación 24). Se presenta en porcentaje.

$$Sp = \frac{\sum TN}{\sum Condition\ Negative\ (TN + FP)} \quad 24$$

Siendo el clasificador ideal uno con una sensibilidad del 100% (todos los píxeles de tejido infartado se identifican como infartados) y 100% específico (todos los píxeles de tejido sano se identifican como sanos). De la misma manera, un accuracy del 100% significaría que los valores medidos son exactamente iguales a las etiquetas establecidas.

De modo que a la hora de optimizar el clasificador generado se va a reparar en estos tres parámetros intentando lograr un equilibrio entre todos ellos.

## Implementación

Para la implementación del clasificador SVM se ha utilizado una herramienta que tiene creada Matlab denominada *ClassificationLearner*. Gracias a esta aplicación que emplea una máquina de aprendizaje supervisado, se pueden crear diferentes clasificadores como árboles de decisión, SVM con diferentes kernels, vecindades, etc. Con esta herramienta es posible llevar a cabo el aprendizaje automático insertando un conjunto conocido de datos de entrada y las etiquetas de estos.

Dicha aplicación da la opción de exportar el modelo generado para así modificarlo y adecuarlo a la aplicación que se plantea y a los datos de entrada. De modo que en primer lugar se ha creado con la herramienta unos modelos preliminares y de ellos se ha escogido aquel que mejor se adapta a los datos de entrada. La elección se ha realizado analizando los valores de accuracy de cada uno de los modelos. Entre todos los que se han probado el más adecuado para los datos de entrada ha sido un clasificador SVM con kernel polinomial de tercer orden, con un accuracy del 96.3%.

Posteriormente, se ha procedido a la selección de los datos que realmente van a ser empleados para el entrenamiento. Debido a la gran diferencia de tamaño de las muestras, la cantidad de píxeles de tejido es 10 veces mayor que los de infartado, antes de realizar la validación cruzada *leave-one-out* se han realizado diferentes pruebas para obtener la relación óptima de entrenamiento entre píxeles de tejido sano y tejido infartado. Estas pruebas se han ejecutado manteniendo en todo momento el 80% de los píxeles pertenecientes al tejido infartado y variando el porcentaje que se coge de los píxeles sanos. En la siguiente imagen (Fig 18) se pueden observar los resultados logrados con los diferentes casos empleados.

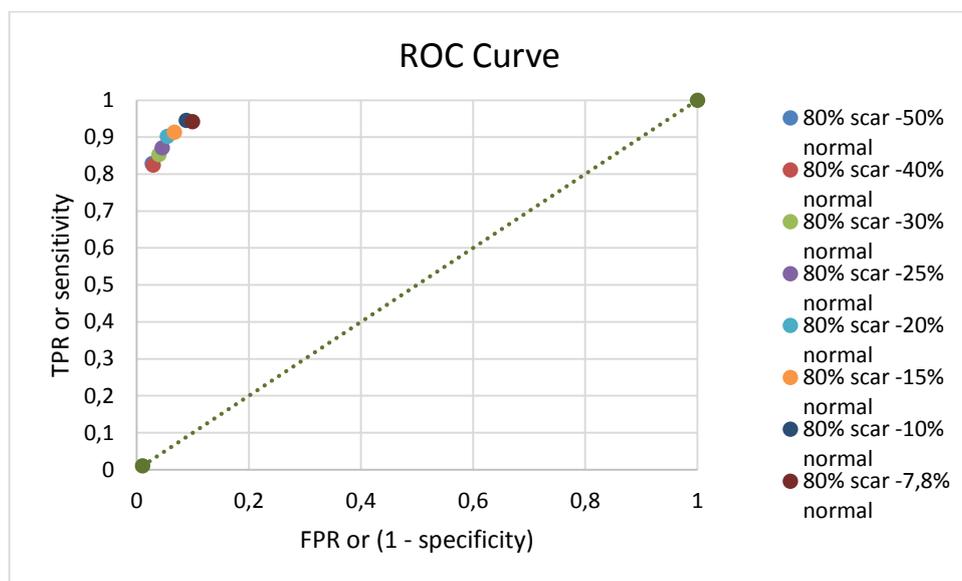


Fig 18. Curva de ROC variando la relación entre píxeles de tejido infartado y sanos.

La curva ROC es una representación gráfica de la sensibilidad frente a la especificidad que tiene un sistema clasificador binario. Dicho gráfico compara los diferentes métodos entre sí, siendo el mejor método posible de predicción el que se coloca en la esquina superior izquierda, en la coordenada (0, 1), contando con el 100% de la sensibilidad y de la especificidad. La línea diagonal que se extiende desde el extremo inferior izquierdo hasta la esquina superior derecha es la línea

de no-discriminación. Esta diagonal divide el espacio ROC en los buenos resultados de clasificación (parte superior) y en los resultados pobres (parte inferior). Consecuentemente, los clasificadores obtenidos pueden ser considerados buenos según la localización de los resultados en la curva ROC.

Ampliando la gráfica en la región en la que se agrupan los resultados (Fig 19) se puede apreciar que conforme se disminuye el porcentaje de píxeles escogidos de tejido sano, aumentando así la proporción de píxeles de tejido dañado, la sensibilidad aumenta pero la especificidad disminuye.

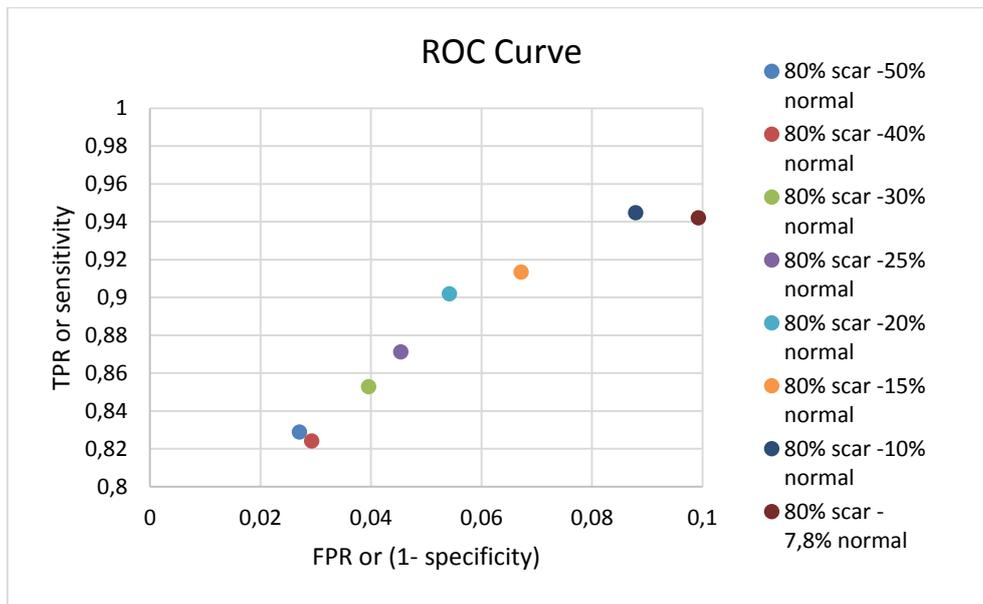
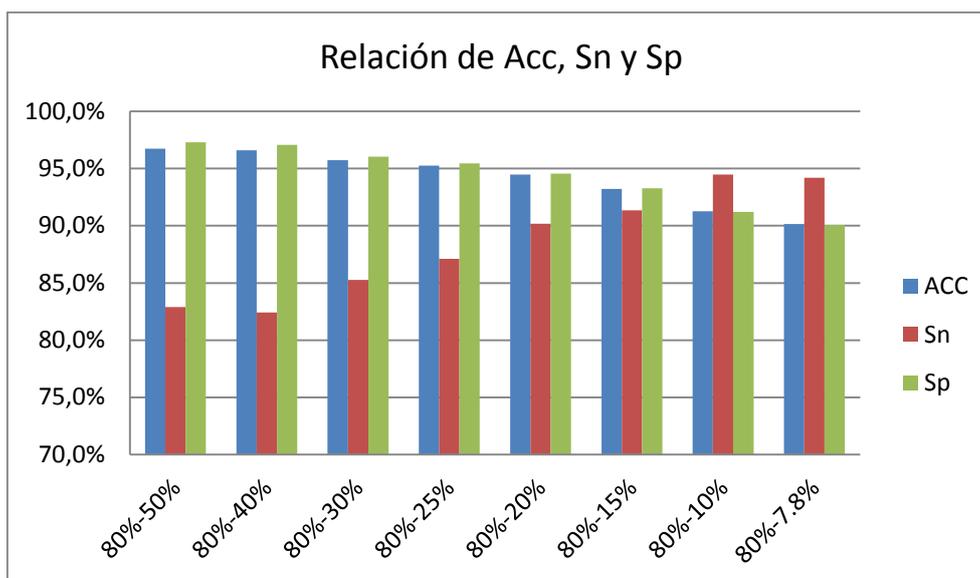


Fig 19. Ampliación de la curva ROC anterior.

En la siguiente gráfica se muestra el cambio que se da en los valores de accuracy ( $Acc$ ), sensibilidad ( $S_n$ ) y especificidad ( $S_p$ ) de cada uno de los casos realizados.



**Fig 20.** Valores de Accuracy, sensibilidad y especificidad de todas las opciones estudiadas.

En ambas figuras (Fig 19 y Fig 20) se puede apreciar que las relaciones que mantienen valores tanto de sensibilidad como de especificidad altos son las de 80%-15% y 80%-10%, obteniendo en ambas relaciones un valor mayor del 90% en accuracy, sensibilidad y especificidad. De estas dos opciones se ha considerado que el clasificador que mejor responde a los datos del estudio es el que utiliza para entrenar el 80% de los píxeles de tejidos infartados y el 15% de los tejidos sanos, con el que se obtiene una exactitud del 93.2%, una sensibilidad del 91.3% y una especificidad del 93.3%. Estos datos han sido conseguidos gracias a las ecuaciones 22, 23 y 24 anteriormente comentadas.

Una vez escogida la relación que se va a emplear para entrenar entre los píxeles sanos y dañados, se ha procedido a la validación cruzada de *leave-one-out* para obtener unos resultados más reales, ya que en este caso el clasificador en el entrenamiento no verá ningún píxel de la imagen que se deje fuera. Además, de las imágenes que son empleadas para entrenar (todas excepto una) se ha entrenado el clasificador con el 100% de los píxeles etiquetados por el especialista como tejido infartado y siguiendo con la relación óptima escogida en el paso anterior, el 18.75% de los píxeles sanos. En la siguiente figura (Fig 21) se muestra de manera más visible la forma en la que se ha realizado la división de los datos para el entrenamiento y el testeo.

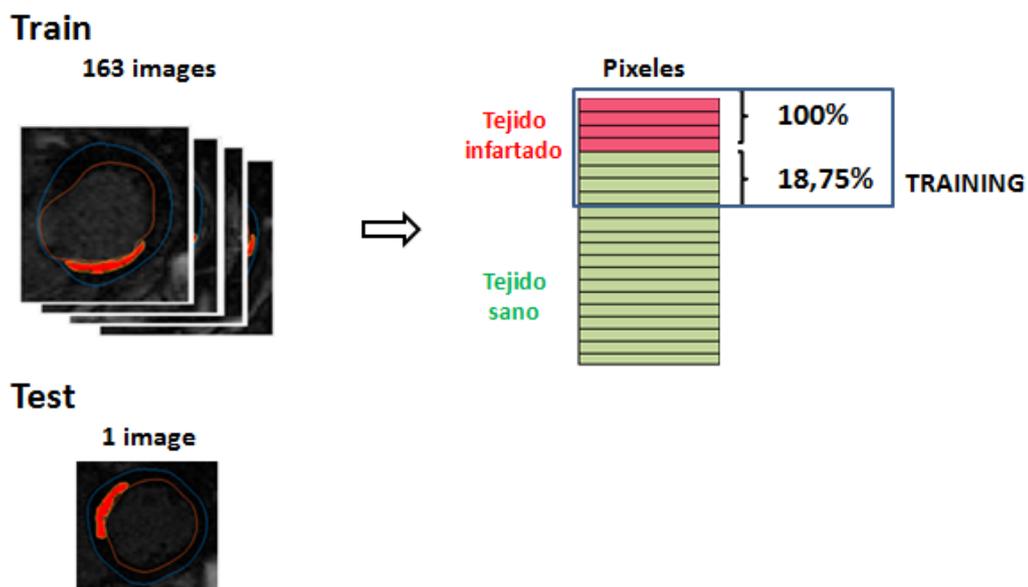


Fig 21. División de los datos para trabajar con el clasificador.

Con el objetivo de tener resultados más realistas, también se ha realizado un proceso de *leave-one-out* eliminando del entrenamiento un usuario entero y empleando este para la evaluación del clasificador. En el siguiente apartado se analizarán los resultados obtenidos con estos clasificadores.



# Capítulo 4

## Resultados preliminares

---

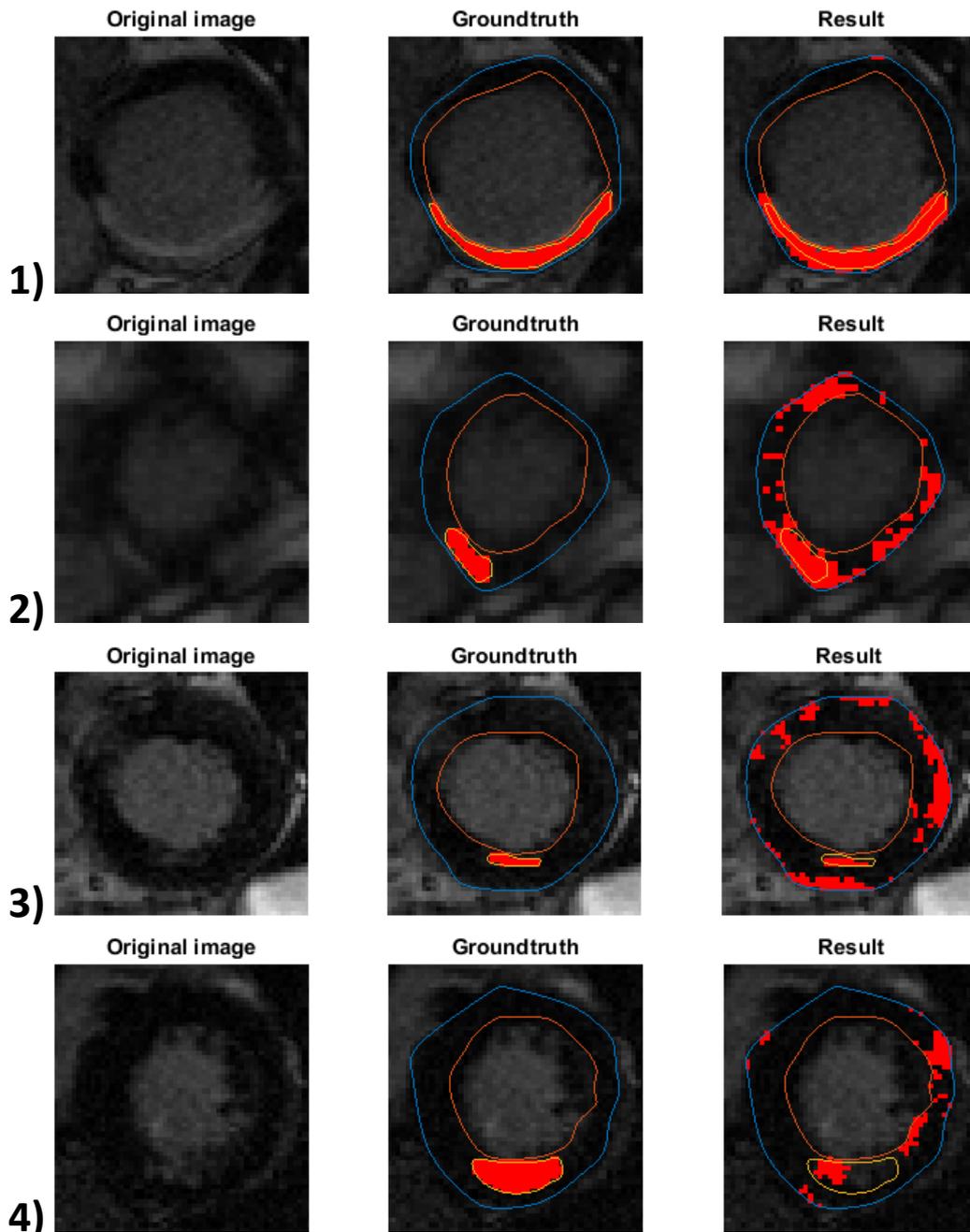
En este capítulo se van a exponer los resultados preliminares obtenidos tras ser testeado el clasificador generado. Como se ha mencionado en el capítulo anterior, tras escoger el porcentaje más adecuado para el conjunto de datos con el que se está trabajando (100% de los píxeles de tejido infartado y el 18.5% de los de tejido sano) se han realizado dos procesos de validación cruzada *leave-one-out*. En la primera prueba se han dejado para testear en cada ronda los píxeles pertenecientes a una imagen, mientras que en la segunda se han dejado los píxeles pertenecientes a un volumen entero, es decir de todas las imágenes de un usuario.

En ambas pruebas se han calculado para cada ronda realizada los valores de accuracy, sensibilidad y especificidad que tiene el clasificador entrenado, obteniendo en la primera prueba un total de 164 resultados (número total de imágenes) y en la segunda prueba 20 (número total de volúmenes).

### **Validación dejando una imagen para testear:**

Tras realizar las 164 rondas se han analizado todas las imágenes resultantes en las que se han coloreado los píxeles clasificados como tejido infartado. A continuación (Fig 22) se muestran cuatro ejemplos diferentes de dichos resultados. En la primera columna de estos se puede

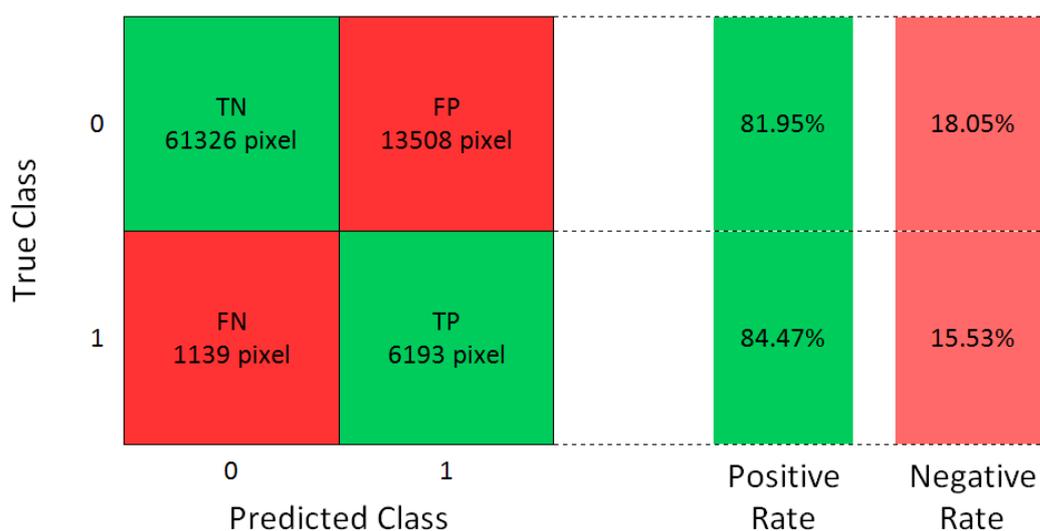
apreciar la imagen original proporcionada por el especialista, ampliada para ver con mayor exactitud el ventrículo izquierdo; en el centro aparece el *groundtruth*, con los píxeles etiquetados como tejido infartado marcados en rojo y finalmente, en la última imagen aparecen coloreados en rojo aquellos píxeles que han sido catalogados según el presente clasificador como tejido dañado.



**Fig 22.** Cuatro ejemplos de los resultados obtenidos empleando la validación cruzada dejando una imagen entera para testear.

En los ejemplos se recogen cuatro casos que se han dado tras realizar la prueba. Tal y como muestra el primero, en muchos casos los píxeles del tejido infartado y del sano se han detectado de manera correcta. Por otro lado, también se dan casos (2 y 3) en los que a pesar de que el área que cubre el tejido infartado haya sido detectada prácticamente en toda totalidad, también hay una gran cantidad de falsos positivos, tejido sano clasificado erróneamente como infartado. Finalmente, el último caso (4) observado se consigue detectar parte del tejido infartado pero no toda la cicatriz marcada, habiendo también tejido sano clasificado erróneamente. Por lo tanto, imágenes con una cantidad elevada de falsos positivos y falsos negativos.

Teniendo en cuenta el total de los píxeles clasificados en las 164 rondas realizadas en esta prueba, en la siguiente matriz de resultados (Fig 23) se puede observar la cantidad de píxeles que han sido clasificados correcta y erróneamente y cuál es el ratio de verdaderos positivos y falsos negativos de estos:



**Fig 23.** Matriz de resultados de la prueba de validación cruzada dejando una imagen entera para testear.

La matriz de resultados presentada muestra, por un lado, en la parte izquierda el modo en el que se han clasificado los píxeles, marcando en el eje vertical las etiquetas reales, las marcadas por el especialista, y en el horizontal las predicciones realizadas. Por otro lado, en la parte derecha aparecen los ratios de verdaderos positivos y falsos negativos, mostrando que del total de los píxeles analizados más del 80% están bien clasificados, tanto los verdaderos positivos (84.47%) como verdaderos negativos (81.95%).

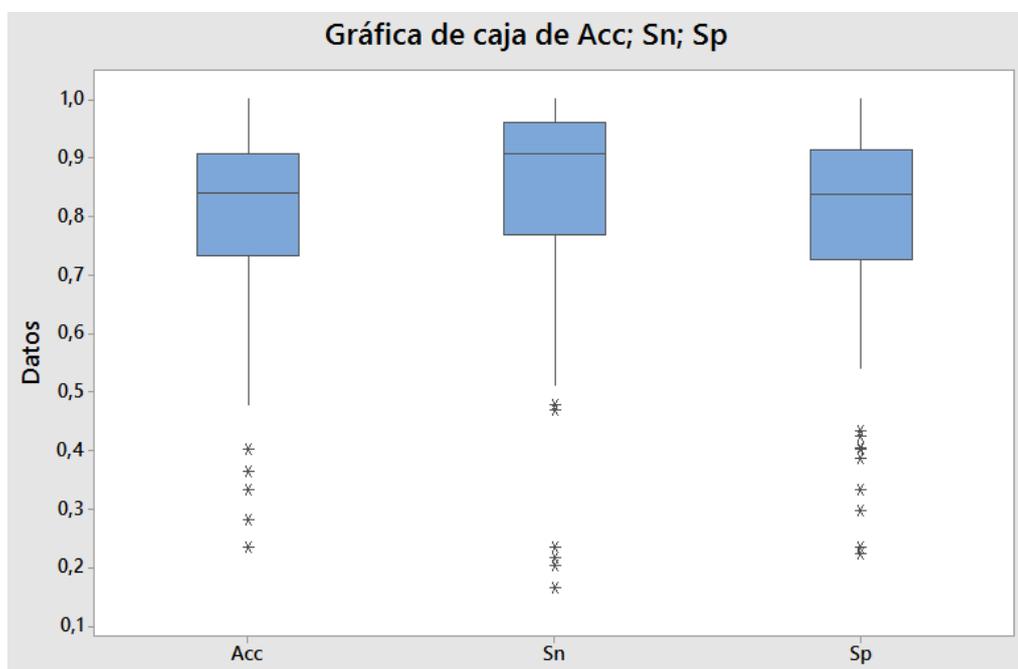
Con los valores presentados anteriormente se ha calculado para cada una de las rondas los valores de accuracy, sensibilidad y especificidad. La Tabla 1 recoge los estadísticos descriptivos obtenidos tras el cálculo de estos:

**Tabla 1.** Estadísticos descriptivos de los resultados obtenidos en la validación cruzada dejando una imagen entera para testear.

Variable	Media	Desv. Est.	Q1	Mediana	Q3
Acc	0.8067	0.1436	0.7321	0.8412	0.9072
Sn	0.8307	0.1902	0.7682	0.9063	0.9599
Sp	0.8004	0.1588	0.7271	0.8376	0.9140

Entre los valores mostrados en la tabla anterior se puede apreciar que en los tres parámetros se tiene una media superior al 80%, es decir, que de media más del 80% de los píxeles están bien clasificados, siendo la sensibilidad el parámetro de mayor media (83.07%). Es necesario tener en cuenta que estos datos cuentan con una variabilidad de entre 15 y 20 puntos.

El resto de los resultados descriptivos se pueden observar de manera más visual en la siguiente representación gráfica de *box-plot*:



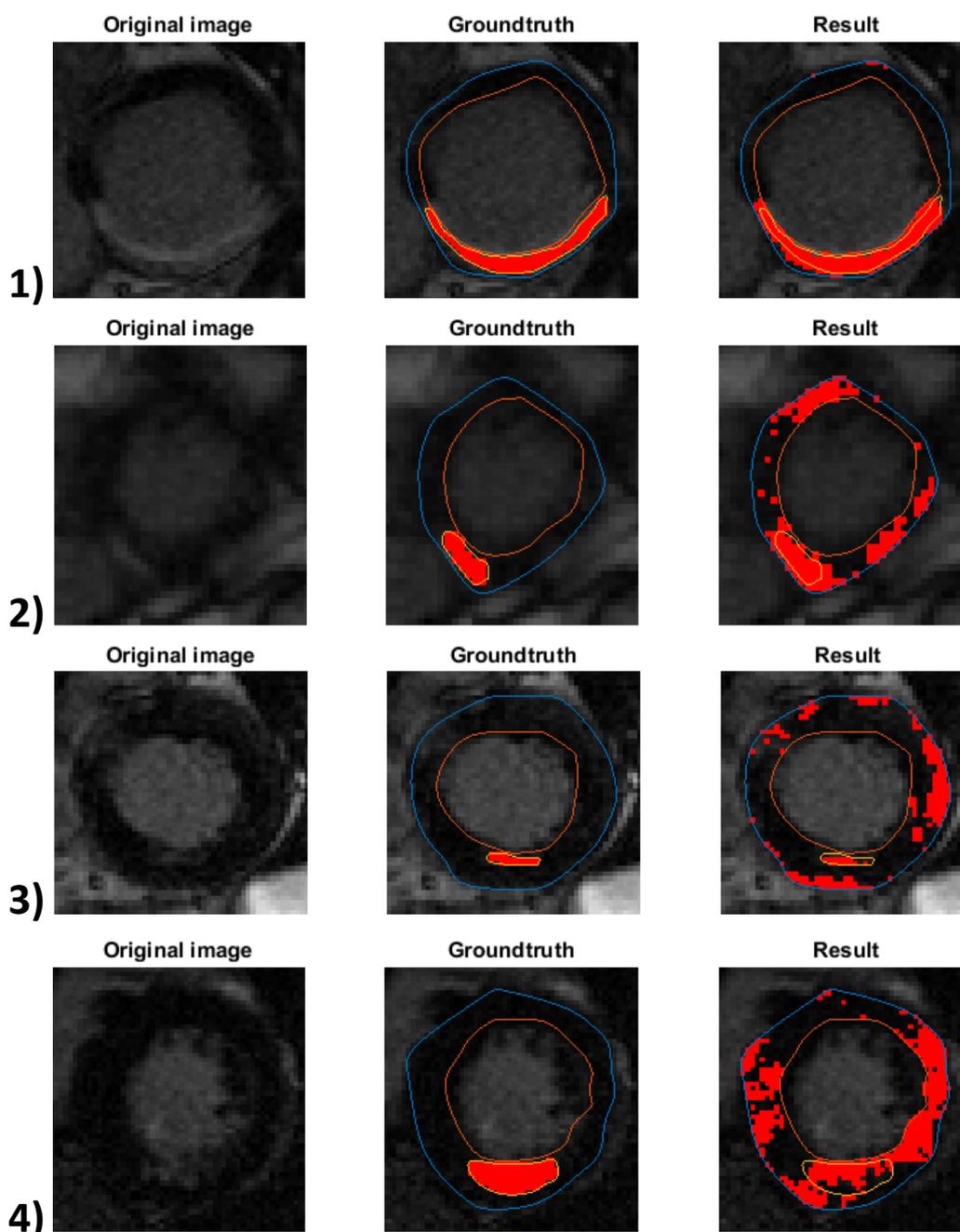
**Fig 24.** Box-plot de los resultados obtenidos en la prueba, dejando imágenes enteras para testear.

Tanto en la (Tabla 1) como en la gráfica (Fig 24) anterior se puede observar que la mediana en los tres parámetros se encuentra por encima del 0.8, es decir, la mitad de las observaciones son mayores que o iguales al 80%, llegando en la sensibilidad hasta un 90.63%. Al mismo tiempo, gracias al primer cuartil se puede apreciar que más del 75% de los datos tienen un valor de accuracy, sensibilidad y especificidad mayor que el 70%. Finalmente, en la gráfica también se puede observar que aparece más de un outlier, en los que el valor de estos parámetros es menor que el 50%.

## Validación dejando un volumen para testear:

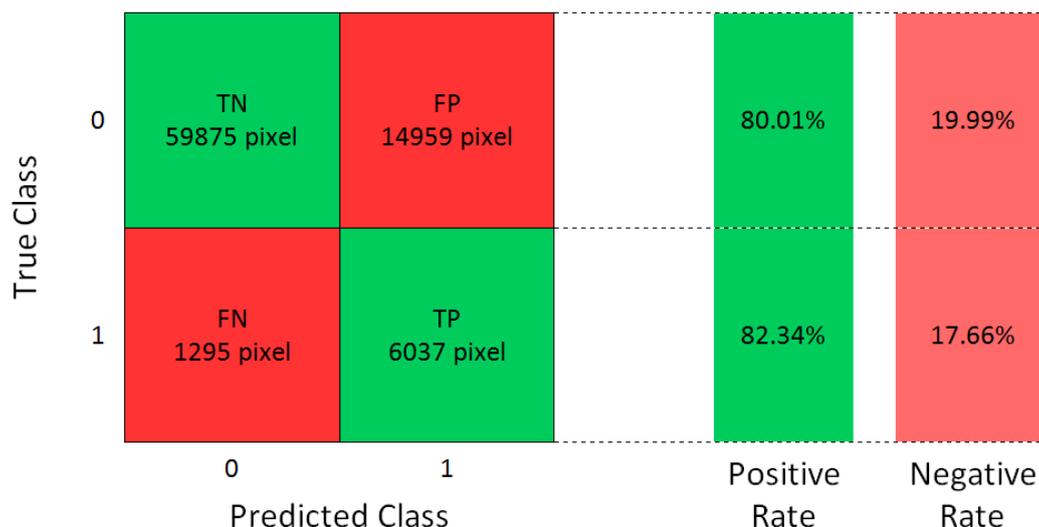
De la misma manera que se ha comentado, se ha repetido el procedimiento anterior excluyendo los estudios enteros a la hora de entrenar. Tras realizar este procedimiento 20 veces se han vuelto a analizar las imágenes resultantes coloreando una vez más los píxeles clasificados como tejido infartado.

A continuación, se muestran (Fig 25) los resultados obtenidos con los mismos ejemplos del apartado anterior. Examinando dichos ejemplos se puede apreciar que excepto en el último caso, los resultados son bastante similares a los del apartado anterior, detectando del mismo modo el área de tejido infartado y aumentando levemente la cantidad de falsos positivos. En la última imagen analizada, en cambio, se visualiza como la cicatriz es detectada con mayor exactitud que en la prueba anterior, pero como a su vez la cantidad de falsos negativos aumenta notoriamente. El empeoramiento de los resultados del clasificador frente a la prueba anterior es lógico, ya que en esta prueba se emplean menos píxeles en el entrenamiento. Además, dentro de los estudios hay imágenes que no varían mucho entre sí, por lo tanto en la prueba anterior el grupo de entrenamiento empleado está compuesto por píxeles de características muy parecidas a los de la imagen destinada a la evaluación. De manera que en el caso anterior el entrenamiento se realiza con píxeles de características similares a las del conjunto de testeo, siendo una visión y evaluación más real la que se plantea en este apartado, excluyendo en el grupo de entrenamiento todos los cortes de un usuario y empleando éste para la evaluación del clasificador.



**Fig 25.** Cuatro ejemplos de los resultados obtenidos empleando una evaluación de validación cruzada dejando un estudio entero para testear.

Tras examinar los resultados visuales del clasificador, en la siguiente matriz (Fig 26) se pueden analizar los resultados cuantitativos de esta prueba. La matriz muestra que, teniendo en cuenta todos los píxeles clasificados en los 20 procesos de entrenamiento realizados (dejando en cada uno un estudio completo fuera para usarlo en el entrenamiento), el 80.01% de los píxeles sanos y el 82.34% de los píxeles de tejido infartado es clasificado correctamente.



**Fig 26.** Matriz de resultados de la prueba de validación cruzada dejando un estudio entero para testear.

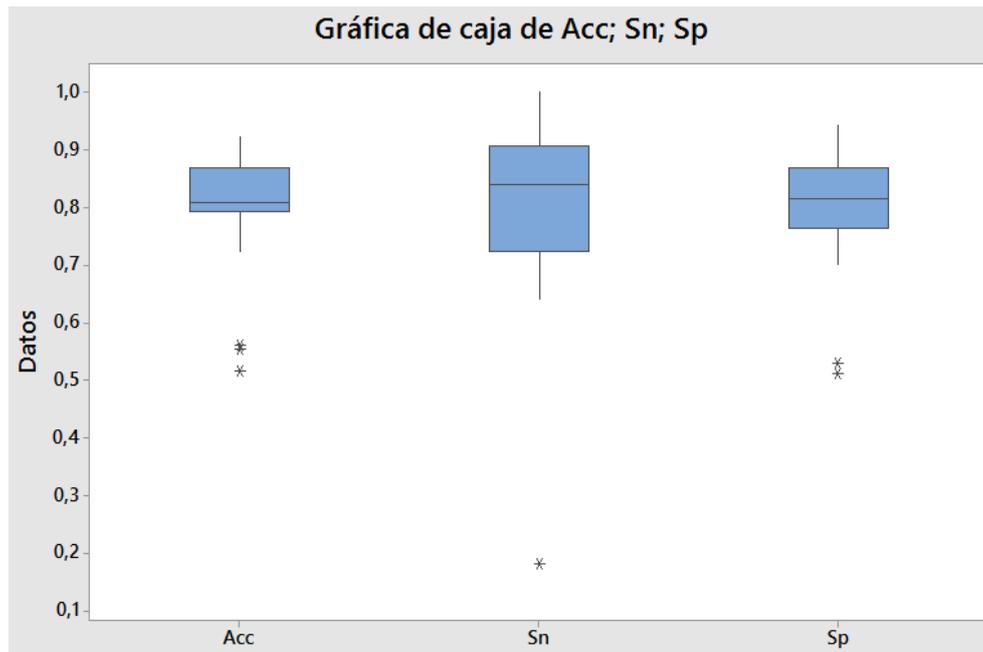
Al igual que en la prueba anterior, en esta también han sido extraídos los resultados de accuracy, sensibilidad y especificidad obtenidos en cada ronda y se han calculado los descriptores estadísticos de ellos. En la Tabla 2 se recogen la media, la desviación estándar, el primer y tercer cuartil y la mediana:

**Tabla 2.** Estadísticos descriptivos de los resultados obtenidos en la validación cruzada dejando un usuario entero para testear.

Variable	Media	Desv. Est.	Q1	Mediana	Q3
<b>Acc</b>	0.7913	0.1170	0.7938	0.8079	0.8692
<b>Sn</b>	0.7989	0.1783	0.7237	0.8394	0.9060
<b>Sp</b>	0.7878	0.1268	0.7626	0.8147	0.8692

Tal y como se ha observado en los ejemplos visuales añadidos, los resultados cuantitativos muestran un ligero empeoramiento de los resultados. Aun así, la media de los tres parámetros sigue siendo cercana al 80%, reduciéndose solo tres puntos en el caso más llamativo, la sensibilidad, que pasa del 83.07% al 79.89%.

El resto de estadísticos se aprecian mejor en la gráfica de caja mostrada a continuación (Fig 27):



**Fig 27.** Blox-pot que recoge los resultados obtenidos en la prueba dejando usuarios completos para testear.

Tanto en la (Tabla 2) como en la gráfica (Fig 27) anterior se puede observar que la mediana de los tres parámetros es mayor que 0.8, y que más del 75% de los datos tienen un porcentaje de accuracy, sensibilidad y especificidad mayor que el 70%. Finalmente, en la gráfica también se aprecia que aparece más de un outlier. La cantidad de valores atípicos es menor en esta prueba debido a que los resultados se recogen por estudios, no por imágenes como en es el caso de la prueba anterior.

## Capítulo 5:

# Optimización y resultados finales

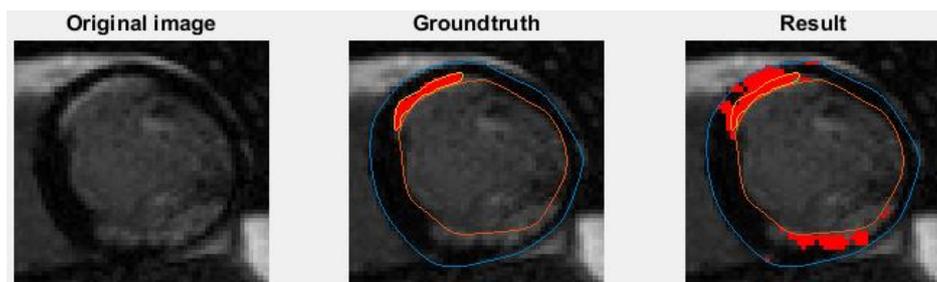
---

En este capítulo se presentan los procesos realizados con el objetivo de optimizar los resultados preliminares presentados en el capítulo anterior, para obtener los resultados finales. Al ser más realista la segunda prueba, la validación cruzada dejando para testear un volumen entero, el post-procesado se ha realizado a las imágenes obtenidas con esta.

Finalmente, en el último apartado se van a comparar los resultados obtenidos tras la optimización con los resultados conseguidos por el especialista empleando una herramienta comercial.

## Eliminación de imágenes inadecuadas

En primer lugar, han sido eliminadas algunas imágenes en las que se había confirmado la presencia de tejidos dañados sin ser etiquetados.



**Fig 28.** Ejemplo de tejido infartado sin etiquetar.

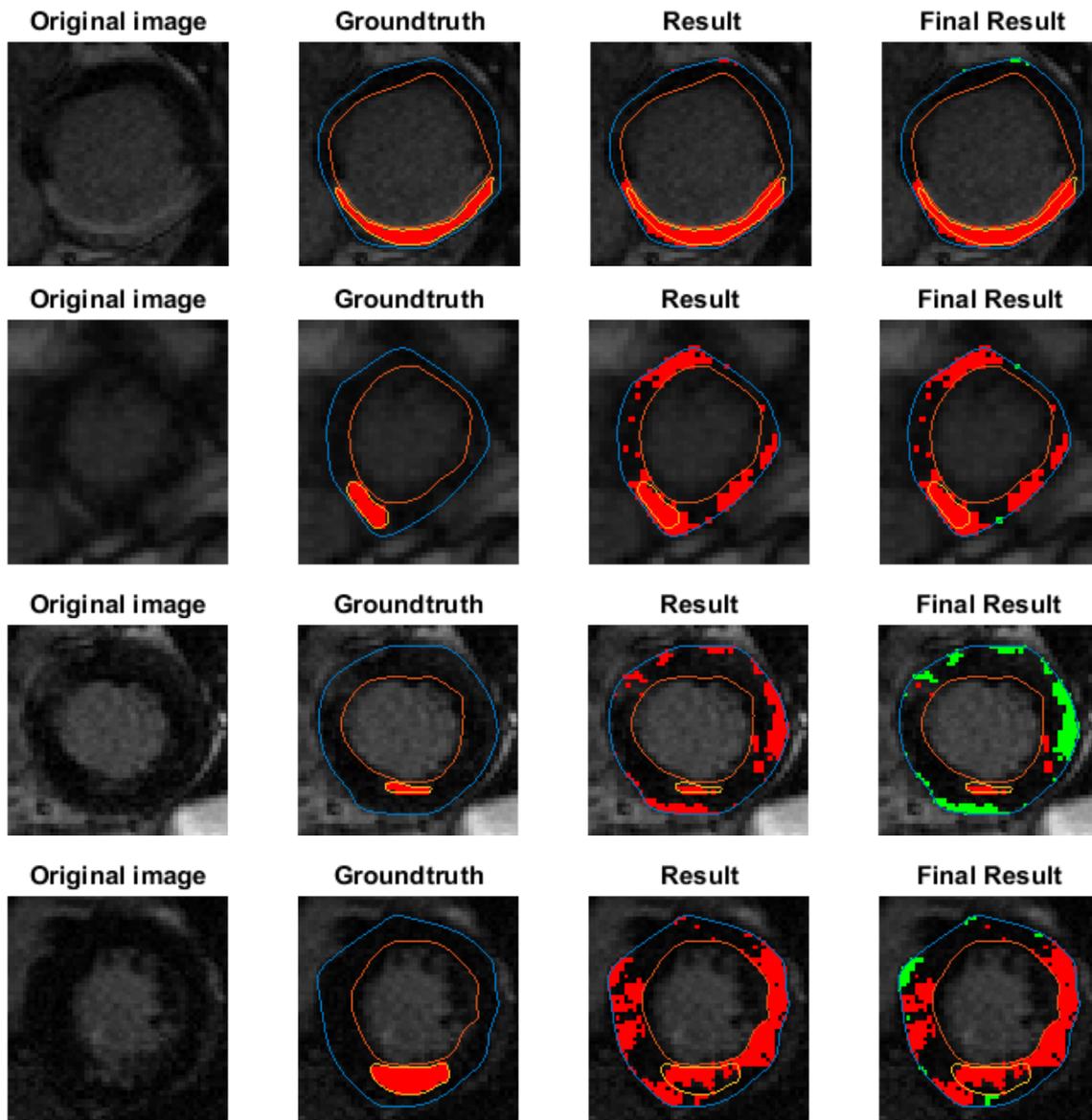
En la figura anterior (Fig 28) se puede apreciar un ejemplo de una de las imágenes que ha sido eliminada. En este caso, tal y como el clasificador generado lo predice, a parte de la escara ya etiquetada en la parte superior de la imagen, también se puede observar en la parte inferior, dentro del miocardio, otro conjunto de píxeles que concuerdan con las características de los píxeles de tejidos dañados. Tras consultar con el especialista las imágenes en las que existía alguna duda, se han excluido del entrenamiento un total de 10 imágenes, evitando así entrenar el clasificador con píxeles mal etiquetados.

## Post-procesado

En segundo lugar, tras analizar todos los resultados del apartado anterior, se ha concluido que es muy difícil aumentar la sensibilidad (el número de píxeles de tejido dañado clasificados correctamente) sin reducir drásticamente la especificidad (número de píxeles de tejido sano clasificados correctamente). Por lo tanto, se ha procedido a buscar un método de post-procesado que sea capaz de eliminar aquellos segmentos clasificados como tejido infartado que en realidad corresponden a tejido sano.

Como ya se ha explicado anteriormente, el deterioro del tejido miocárdico se genera del borde del endocardio hacia fuera. Tras analizar cada una de las imágenes se ha comprobado que, en efecto, todas las cicatrices se encuentran pegadas al endocardio, y que en caso de que toquen el epicardio es porque la cicatriz atraviesa de lado a lado el miocardio. Por consiguiente, se entiende que todos aquellos segmentos clasificados como tejido infartado que están pegados solamente a la parte exterior del miocardio no son cicatriz y han sido eliminados.

A continuación (Fig 29) se pueden contemplar los resultados obtenidos tras dicho post-procesado en los tres ejemplos anteriormente comentados, a los resultados obtenidos anteriormente se les ha añadido una cuarta columna en la que aparece el resultado final, coloreando de verde los píxeles que en un principio se habían considerado tejido dañado y han sido eliminados tras el post-procesado, y de rojo los píxeles resultantes.



**Fig 29.** Cuatro ejemplos de los resultados obtenidos tras el post-procesado.

En estos ejemplos se puede apreciar que gracias al post-procesado en muchos casos se reduce una gran cantidad de falsos positivos. Aun así, sigue habiendo casos en los que no se pueden eliminar y en los que sería necesario un post-procesado más complejo. En el Anexo I se pueden visualizar los resultados de todas las imágenes analizadas.

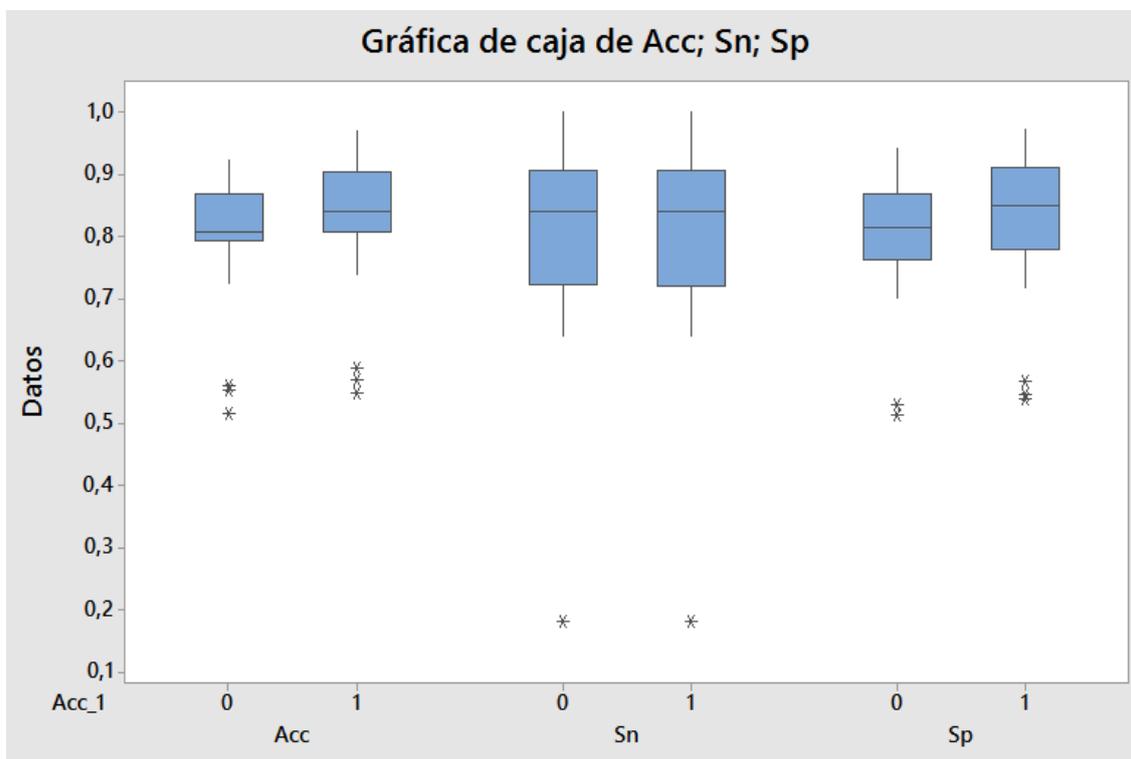
Con el objetivo de ver cuantitativamente cómo mejoran los resultados tras este post-procesado, se han extraído, como en los resultados preliminares los valores de accuracy, sensibilidad y especificidad. En la siguiente Tabla 3 se recogen los descriptores estadísticos de dichos parámetros:

**Tabla 3.** Estadísticos descriptivos de los resultados obtenidos tras el post-procesado.

Variable	Media	Desv. Est.	Q1	Mediana	Q3
Acc	0.8218	0.1228	0.8067	0.8405	0.9043
Sn	0.7987	0.1784	0.7209	0.8394	0.9060
Sp	0.8202	0.1327	0.7786	0.8489	0.9100

Gracias al post-procesado realizado, los estadísticos que describen el accuracy y la especificidad han mejorado, aumentando por ejemplo en la media algo más de tres puntos en ambos parámetros. Por otro lado, con este proceso los resultados de sensibilidad no empeoran a los ya obtenidos, por lo que se puede concluir que se ha conseguido cumplir el objetivo de mejorar el clasificador sin empeorar ninguno de los tres parámetros.

A continuación (Fig 30) se puede observar una gráfica de box-plot que contiene los resultados preliminares (correspondientes a la etiqueta 0) de la validación cruzada dejando estudios enteros para entrenar y los resultados finales tras el post-procesado (correspondientes a la etiqueta 1) de las imágenes resultantes.



**Fig 30.** Box-plot comparativo de los resultados preliminares, marcados con 0, y los post-procesados, representados con el 1.

En el box-plot se observa claramente como tanto en accuracy como en la especificidad gracias al post-procesado se ha mejorado el clasificador, ya que las distribuciones de los resultados que

corresponden a los de post-procesado (1) se encuentran más arriba que los resultados preliminares (0). También se aprecia como en el caso de la sensibilidad no hay ningún cambio tras el post-procesado.

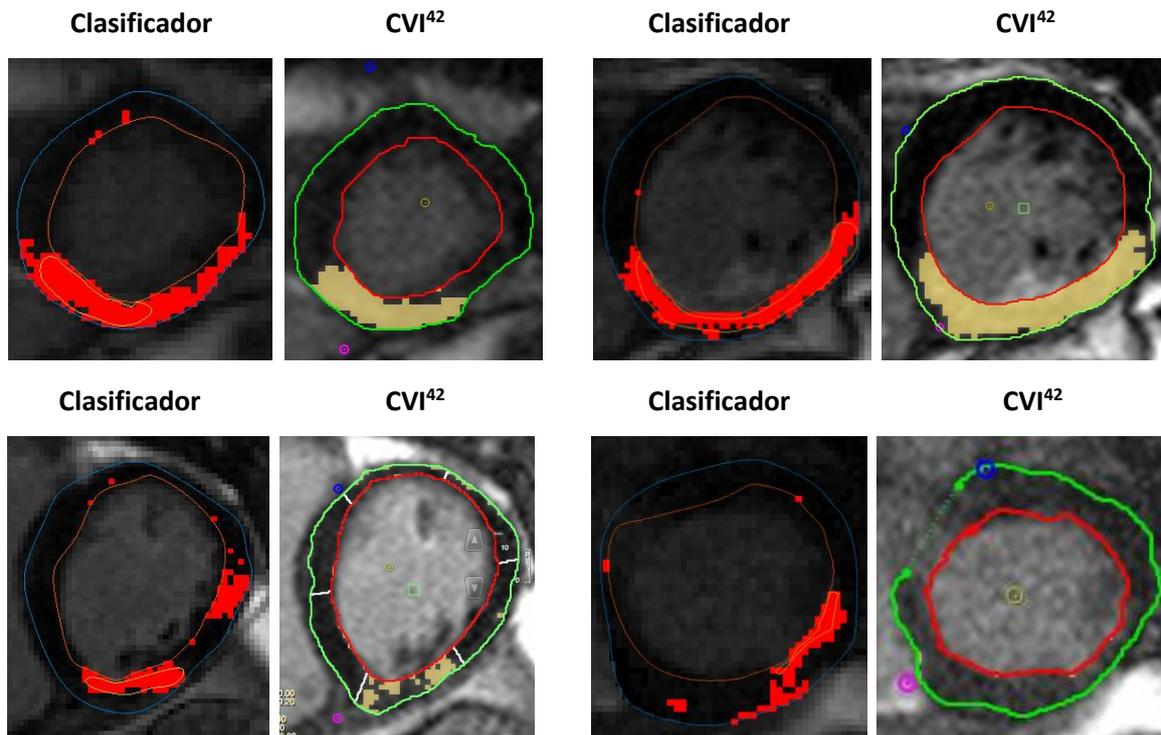
## Comparativa de resultados finales con resultados de CVI<sup>42</sup>

Siete de los volúmenes empleados en la generación del clasificador han sido analizados con el programa comercial que emplea el especialista: CVI<sup>42</sup>, desarrollado por *Circle Cardiovascular Imaging Inc.*, se trata de un software para la evaluación de imágenes cardíacas de RM y TC. Entre otras funciones, esta herramienta permite realizar la caracterización del tejido miocárdico, permitiendo evaluar así la extensión y la distribución regional de una lesión (cicatriz, edema, inflamación) del corazón [16].

El CVI<sup>42</sup> efectúa la detección de tejido infartado mediante una umbralización basada en el nivel de gris de los píxeles, para ello la herramienta ofrece varios métodos de umbralización, entre las que se encuentran una función manual en la que el usuario selecciona el umbral que considera más adecuado y una función que define el umbral como la media más tres veces la desviación estándar.

En el análisis de los siete volúmenes, el especialista ha empleado la función manual, ya que actualmente es la que se emplea en la práctica clínica. En esta función, el especialista para comenzar tiene que marcar en las imágenes elementos concretos de la anatomía del corazón, como son la unión ventricular, el ápex, etc. También debe delimitar el miocardio. Posteriormente, estudia el efecto de diferentes umbrales en las imágenes y finalmente establece el umbral más adecuado para la detección de los segmentos que corresponden a tejido infartado.

A continuación (Fig 31) se puede observar la comparación de cuatro resultados visuales obtenidos mediante la herramienta CVI<sup>42</sup> con la función manual y el clasificador creado en este trabajo:



**Fig 31.** Comparación cualitativa de resultados. Izquierda: resultados con el clasificador; Derecha: resultados de CVI<sup>42</sup>.

En los ejemplos se observa como la herramienta creada detecta más píxeles como tejido infartado que el software CVI<sup>42</sup>, lo cual por un lado genera una mayor cantidad de falsos positivos, pero por otro es capaz de detectar mayor número de píxeles de tejido infartado, consiguiendo así en el último ejemplo añadido detectar una cicatriz que el CVI<sup>42</sup> no ha sido capaz de detectar.

Por otro lado, los resultados de CVI<sup>42</sup> mostrados han sido obtenidos mediante la umbralización manual, lo que supone que se trata de una función dependiente del especialista, ya que este ha sido él, quien ha fijado el umbral que considera adecuado para la detección. Además, dicho análisis le supone al especialista alrededor de 30 minutos por cada estudio. Sin embargo, mediante la herramienta creada en este trabajo, una vez se han marcado los bordes del miocardio, el proceso de detección es totalmente automático y el tiempo necesario para realizar el análisis de un estudio, desde la extracción de características hasta el resultado final, está en torno a los 5 minutos. Por lo tanto, el software consigue reducir claramente la carga de trabajo al especialista.

Con el objetivo de comparar los resultados de los dos clasificadores cuantitativamente, se ha extraído para cada corte el porcentaje de tejido infartado detectado con cada una de las herramientas, teniendo en cuenta todo el miocardio. En la siguiente tabla (Tabla 4) se recogen dichos porcentajes, junto al porcentaje que suponen las marcas realizadas por el especialista, el *groundtruth*. A la hora de analizar los resultados mostrados, es necesario tener en cuenta que no se ha marcado el miocardio de la misma manera en los dos clasificadores, por lo que el 100% del miocardio no es el mismo en los resultados de una herramienta y de la otra.

**Tabla 4.** Comparativa de resultados cuantitativos entre el clasificador generado y el CVI<sup>42</sup>.

Volumen	% Cicatriz marcada	%Cicatriz clasificador	%Cicatriz CVI <sup>42</sup>
Caso 1 (Im 043 – Im 048)	14.53%	14.13%	10.10%
	20.86%	25.40%	15.60%
	19.72%	26.60%	22.20%
	23.36%	32.67%	24.30%
	22.98%	37.02%	18.00%
Caso 2 (Im 034 – Im 042)	10.17%	23.31%	12.90%
	21.58%	42.09%	30.80%
	24.06%	36.19%	32.90%
	24.67%	30.47%	26.70%
	20.55%	34.26%	23.20%
	23.03%	38.49%	29.40%
	26.10%	43.01%	24.80%
Caso3 (Im 025 – Im 033)	20.33%	42.09%	24.30%
	16.89%	40.60%	21.70%
	9.92%	44.84%	0.20%
	0.00%	21.40%	2.40%
	18.61%	23.31%	30.70%
	13.11%	18.03%	23.70%
	16.69%	20.13%	25.10%
	24.54%	32.72%	34.80%
Caso4 (Im 018 – Im 024)	24.49%	34.18%	30.40%
	22.82%	33.02%	0.00%
	25.24%	31.55%	24.40%
	14.67%	21.67%	11.90%
	7.82%	59.68%	14.00%
Caso5 (Im 049 – Im 057)	10.06%	28.21%	9.40%
	7.88%	50.33%	5.60%
	13.20%	41.47%	4.80%
	8.67%	39.52%	4.60%
	8.17%	41.50%	5.80%
Caso5 (Im 049 – Im 057)	2.06%	27.21%	1.70%
	0.00%	21.21%	0.10%
	6.74%	14.42%	3.60%
	3.97%	5.58%	2.40%
	3.54%	5.44%	3.70%

	4.53%	6.65%	2.50%
	1.70%	9.28%	2.20%
	0.00%	9.27%	1.20%
	0.00%	14.72%	0.40%
Caso 6 (Im 001 – Im 008)	0.00%	18.65%	5.00%
	16.76%	24.78%	9.80%
	12.84%	25.68%	6.00%
	7.76%	19.86%	11.40%
	0.00%	17.65%	1.10%
	0.00%	19.95%	1.30%
	0.00%	6.56%	0.10%
	0.00%	1.79%	1.00%
Caso7 (Im 148 – Im 157)	0.00%	8.81%	0.30%
	3.40%	14.56%	1.30%
	1.83%	7.97%	1.60%
	0.00%	0.80%	1.20%
	0.00%	0.20%	1.10%
	0.00%	0.51%	0.70%
	0.00%	8.17%	3.70%
	0.00%	18.23%	3.30%
	0.00%	34.30%	4.40%
0.00%	19.53%	6.20%	

Se ha calculado el promedio de las diferencias entre los resultados del clasificador generado y del CVI<sup>42</sup> con el fin de poder conocer cuál de los dos clasificadores se acercan cuantitativamente más a las marcas realizadas el especialista. En el caso del detector generado en este proyecto, se ha medido una desviación de 13.87 puntos, mientras que en la herramienta comercial ha sido de 4.01 puntos. A pesar de ser una medida orientativa, porque el total del miocardio varía de una prueba a otra, era de esperar que la desviación en el caso de la herramienta comercial sea menor, ya que ha sido el mismo especialista quién ha escogido el valor del umbral y ha dado el visto bueno a los resultados.

En cuanto a la desviación obtenida con la herramienta desarrollada en este proyecto se considera aceptable para ser una versión inicial del software con margen de mejora.

# Capítulo 6:

## Conclusiones y líneas futuras

---

### Conclusiones

A lo largo de este proyecto se ha desarrollado una herramienta para detectar tejido infartado en imágenes cardíacas de resonancia magnética obtenidas mediante la técnica RTG por RMC. Con esta finalidad se ha implementado un proceso que incluye una etapa de extracción de características, otra reducción de dimensionalidad y finalmente, una propuesta de clasificador y el post-procesado de los resultados. Todo ello se ha realizado en MATLAB.

Una vez extraídas y escogidas las características más significativas y generado un primer modelo de clasificador, uno de los pasos más importantes ha sido la elección de la relación entre píxeles de tejido infartado y sano para el entrenamiento del clasificador, ya que dependiendo de esta relación los resultados varían mucho entre sí. Por ello se han realizado diferentes pruebas cambiando dicha relación. La decisión de escoger la relación adecuada se ha realizado teniendo en cuenta tanto el accuracy como la sensibilidad y la especificidad obtenidas con el modelo. De manera que, entre los resultados estudiados en este proceso se ha entendido como la relación óptima, la que recoge el 80% de los píxeles de cicatriz y el 15% de los píxeles sanos, obteniendo con este modelo un accuracy de 93.2%, una sensibilidad de 91.3% y una especificidad de 93.3%.

Tras escoger los porcentajes de entrenamiento adecuados se han realizado dos pruebas de validación cruzada, en las que se han excluido del entrenamiento una imagen en una y un volumen entero en la otra. El grupo de entrenamiento de ambas pruebas se ha generado teniendo en cuenta la relación óptima descrita, extrapolada para elegir el 100% de los píxeles de tejido infartado y por tanto el 18.5% de píxeles sanos. De este modo, en ambas pruebas se han conseguido valores medios de accuracy, sensibilidad y especificidad de en torno al 80%, siendo levemente peores en la segunda prueba en la que se dejan para testear estudios completos. Dicho empeoro es lógico, ya que en esta prueba se emplean menos píxeles en el proceso de entrenamiento. Además, dentro de los estudios hay imágenes que no varían mucho entre sí, por lo que en la primera prueba el entrenamiento se realiza con píxeles de características similares a las del conjunto de testeo, haciendo que esta prueba sea menos realista.

Puesto que la segunda prueba se acerca más a la realidad, se ha procedido a la optimización de dicho método mediante un post-procesado, eliminando los segmentos clasificados como tejido infartado que se encuentran pegados solamente a la parte exterior del miocardio. De esta manera, se ha conseguido mejorar los resultados de accuracy (82.18%, mejora de 3.05 puntos) y especificidad (82.02%, mejora de 3.24 puntos), manteniendo el valor de sensibilidad (79.87%) intacto.

Finalmente, se ha calculado para cada imagen el porcentaje de los píxeles pertenecientes al miocardio que se han clasificado como tejido infartado y dichos porcentajes han sido comparados con los resultados obtenidos por la herramienta comercial que emplea el especialista, CVI<sup>42</sup>. La comparación se ha realizado calculando el promedio de las diferencias entre los porcentajes estimados por los algoritmos y el porcentaje etiquetado por el especialista como cicatriz. De este modo, se ha observado que cuantitativamente se acercan más a las marcas del especialista los resultados obtenidos con la herramienta comercial, con una desviación de 4.01 puntos, que los de la herramienta desarrollada, desviación de 13.87 puntos. Estos resultados se consideran lógicos, puesto que se emplea una función manual en la que es el mismo especialista quién escoge el valor del umbral adecuado y da el visto bueno a los resultados. Aun así, se considera aceptable la desviación obtenida con la herramienta desarrollada.

Resumiendo, se consideran muy positivos todos los resultados obtenidos mediante la herramienta desarrollada en este proyecto, ya que es capaz de detectar correctamente más del 80% de los píxeles, de proceder automáticamente una vez se han marcado los bordes del

miocardio y de reducir notoriamente el tiempo necesario para la realización del proceso (de 30 minutos a 5), consiguiendo de esta forma reducir considerablemente la carga de trabajo del especialista. Además, debe tenerse en cuenta que en este trabajo se presenta la versión inicial de la herramienta con margen de mejora.

## Líneas futuras de investigación

A continuación se presentan algunas líneas de investigación que pueden ser objeto de interés relacionadas al trabajo presentado.

En primer lugar, dado que se ha desarrollado la herramienta mediante aprendizaje automático, se considera que serían necesarios más volúmenes para conseguir generar un detector más preciso y poder realizar un posterior examen más exhaustivo. Además, puede ser interesante emplear el clasificador generado con la base de datos que proporciona Karim et al. en su artículo [2] y evaluarlo con el marco de evaluación que propone.

En segundo lugar, puesto que uno de los objetivos principales de este proyecto es la reducción de la carga de trabajo del especialista, sería de gran utilidad desarrollar un sistema totalmente automático, en el que el especialista no tenga la necesidad de delimitar los bordes del miocardio. Este proceso es posible de realizar con las herramientas *Active Appearance Model (AAM)* y *Active Shape Model (ASM)*, algoritmos de visión computarizada que tras ser entrenados con imágenes similares son capaces de localizar un objeto en una imagen teniendo en cuenta la apariencia y la forma de este.

Respecto a la mejora en la precisión del clasificador, sería conveniente completar el post-procesado de manera que la herramienta sea capaz de clasificar como tejido infartado los huecos que aparecen sin detectar y están situados entre segmentos de tejido infartado. De esta manera, se podría mejorar la sensibilidad de la herramienta. Incluso podría realizarse el mismo post-procesado teniendo en cuenta los cortes anterior y posterior de la imagen, ya que el tejido dañado debe aparecer al menos en dos cortes.

Dados los resultados positivos obtenidos en este trabajo y la cantidad de líneas futuras posibles que proporcionar un amplio margen de mejora, el desarrollo futuro de la herramienta generada resulta prometedor y de gran interés.



# Bibliografía

---

- [1] A. Fernández-Ortiz, «Capítulo 28. Qué es el infarto agudo de miocardio,» de *Libro de la salud cardiovascular*, F. BBVA, Ed., Bilbao, Editorial Nerea, S. A., 2009, pp. 259-268.
- [2] R. Karim, P. Bhagirath, P. Claus, R. J. Housden, Z. Chen, Z. Karimaghaloo, H.-M. Sohn, L. Lara Rodríguez, S. Vera, X. Albà, A. Hennemuth, H.-O. Peitgen, T. Arbel, M. A. González Ballester, A. F. Frangi, M. Götte, R. Razavi, T. Schaeffter y K. Rhode, «Evaluation of state-of-the-art segmentation algorithms for left ventricle infarct from late Gadolinium enhancement MR images,» *Medical Image Analysis*, nº 30, pp. 95-107, 2016.
- [3] M. A. Pietrani, D. Pérez de Arenaza, N. Cóccharo, A. Cagide, J. G. Krauss y R. García Mónaco, «La extensión del infarto de miocardio medido por resonancia magnética con realce tardío de gadolinio se correlaciona con el deterioro de la función del ventrículo,» *Hospital Italiano de Buenos Aires*, vol. 27, nº 1, pp. 4-6, 2007.
- [4] R. J. Kim, D. S. Fieno, T. B. Parrish, K. Harris, E.-L. Chen, O. Simonetti, J. Bundy, J. P. Finn, F. J. Klocke y R. M. Judd, «Relationship of MRI Delayed Contrast Enhancement to Irreversible Injury, Infarct Age, and Contractile Function,» *Circulation*, vol. 100, nº 19, pp. 1992-2002, 1999.
- [5] M. Breuwer, I. Paetsch, E. Nagel, R. Muthupillai, S. Flamm, S. Plein y J. Ridgway, «The detection of normal, ischemic and infarcted myocardial tissue using MRI,» *Int. Congress Ser.*, vol. 30, nº 19, pp. 2360-2367, 2009.

- [6] L. Amado, B. Gerber, S. Gupta, D. Rettmann, G. Szarf, R. Shock, K. Nasir, D. Kraitchmman y J. Lima, «Accurate and objective infarct sizing by contrast enhanced magnetic resonance imaging in a canine myocardial infarction model.,» *J. Am. Coll. Cardiol.*, vol. 44, nº 12, pp. 2383-2389, 2004.
- [7] V. Positano, A. Pingitore, A. Giorgetti, B. Favilli, M. F. Santarelli, L. Landini, P. Marzullo y M. Lombardi, «A fast and effective method to assess myocardial necrosis by means of contrast magnetic resonance imaging,» *J. Cardiovas. Magn. Reson.* , vol. 7, nº 2, pp. 487-494, 2005.
- [8] A. Hennemuth, A. Seeger, O. Friman, S. Miller, B. Klumpp, S. Oeltze y H.-O. Peitgen, «A comprehensive approach to the analysis of contrast enhanced cardiac MR images,» *IEEE Trans. Med. Imag.*, vol. 27, nº 11, pp. 1592-1610, 2008.
- [9] T. O'Donnell, N. Xu, R. Setser y R. D. White, «Semi-Automatic Segmentation of Non-Viable Cardiac Tissue Using Cine and Delayed Enhancement Magnetic Resonance Images,» *Medical Imaging, A. V. Clough and A. A. Amini, Eds. SPIE*, vol. 5031, nº 1, pp. 242-251, 2003.
- [10] L. P. Kotu, K. Engan, K. Skretting, F. Måløy, S. Ørn, L. Woie y T. Eftestøl, «Probability mapping of scarred myocardium using texture and intensity features in CMR images,» *BioMedical Engineering OnLine*, vol. 12, nº 91, pp. 1-19, 2013.
- [11] R. Morisi, B. Donini, N. Lanconelli, J. Rosengarden, J. Morgan, S. Harden y N. Curzen, «Semi-automated scar detection in delayed enhanced cardiac magnetic resonance images,» *International Journal of Modern Physics, World Scientific*, vol. 26, nº 1, pp. 1550011\_1-1550011\_17 , 2015.
- [12] L. Shapiro y G. Stockman, «Texture,» de *Computer Vision*, 2000, pp. 235-247.
- [13] M. Presutti, «La matriz de co-ocurrencia en la clasificación multispectral: tutorial para la enseñanza de medidas texturales en cursos de grado universitario,» São Leopoldo, RS, Brasil, 2004.
- [14] Z. R. Smith y C. S. Wells, *Central Limit Theorem and Sample Size*, University of Massachusetts Amherst. 31.
- [15] E. J. C. Suárez, «Tutorial sobre Máquinas de Vectores Soporte (SVM),» pp. 1-25, 2014.

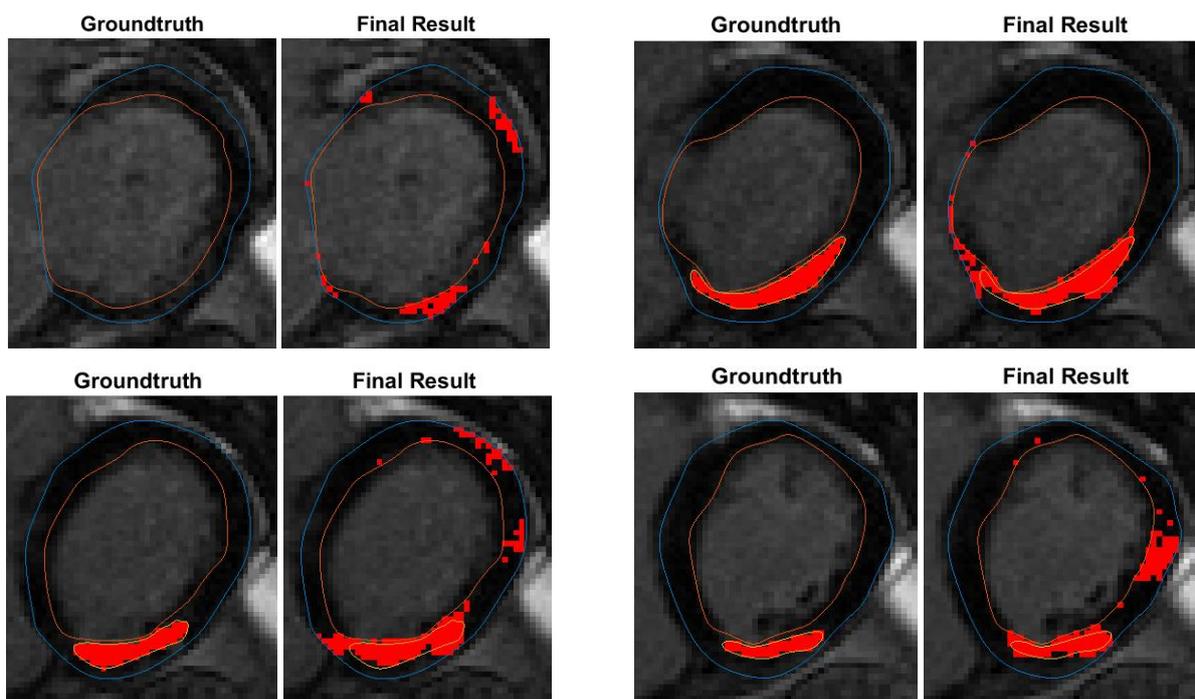
- [16] «Circle Cardiovascular Imaging,» Circle Cardiovascular Imaging, 2016. [En línea]. Available: <http://www.circlecvi.com/products.php#filter=.cvi42>. [Último acceso: 15 Abril 2017].
- [17] M. Rajchl, J. Yuan, J. A. White, E. Ukwatta, J. Stirrat, C. M. S. Nambakhsh, F. P. Li y T. M. Peters, «Interactive Hierarchical-Flow segmentation of Scar Tissue From Late-Enhancement Cardiac MR Images,» *IEEE Transaction on Medical Imaging*, vol. 33, nº 1, pp. 159-172, 2014.

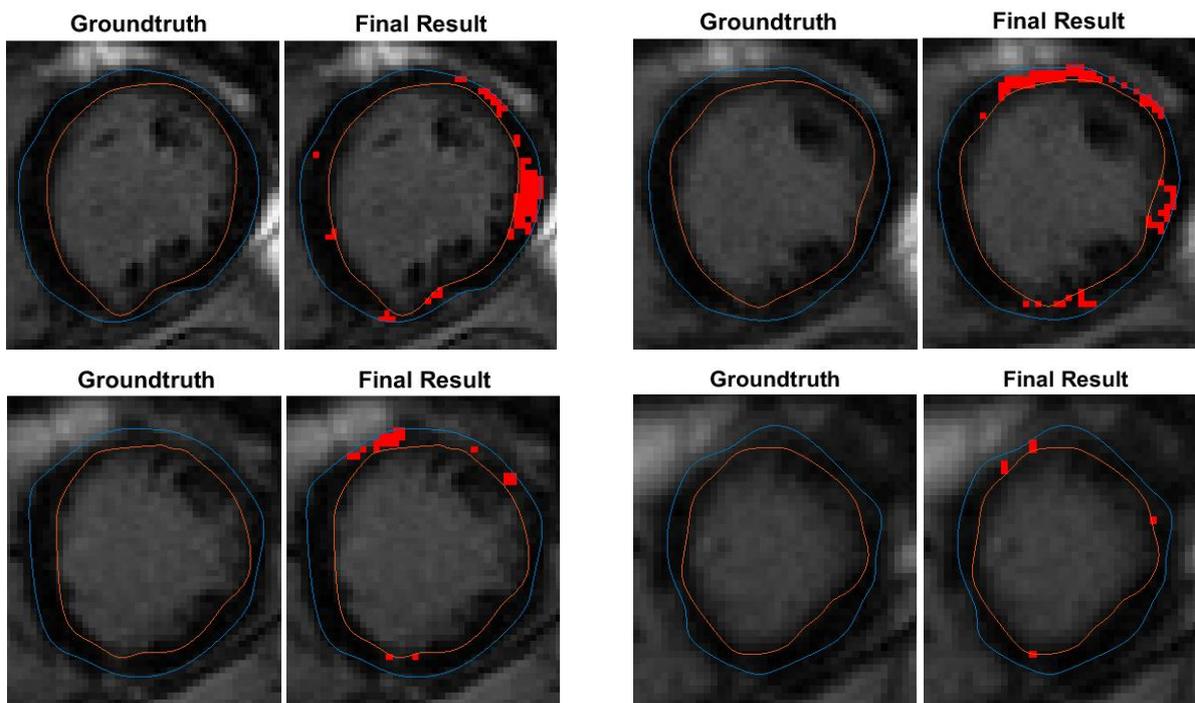


# Anexo I

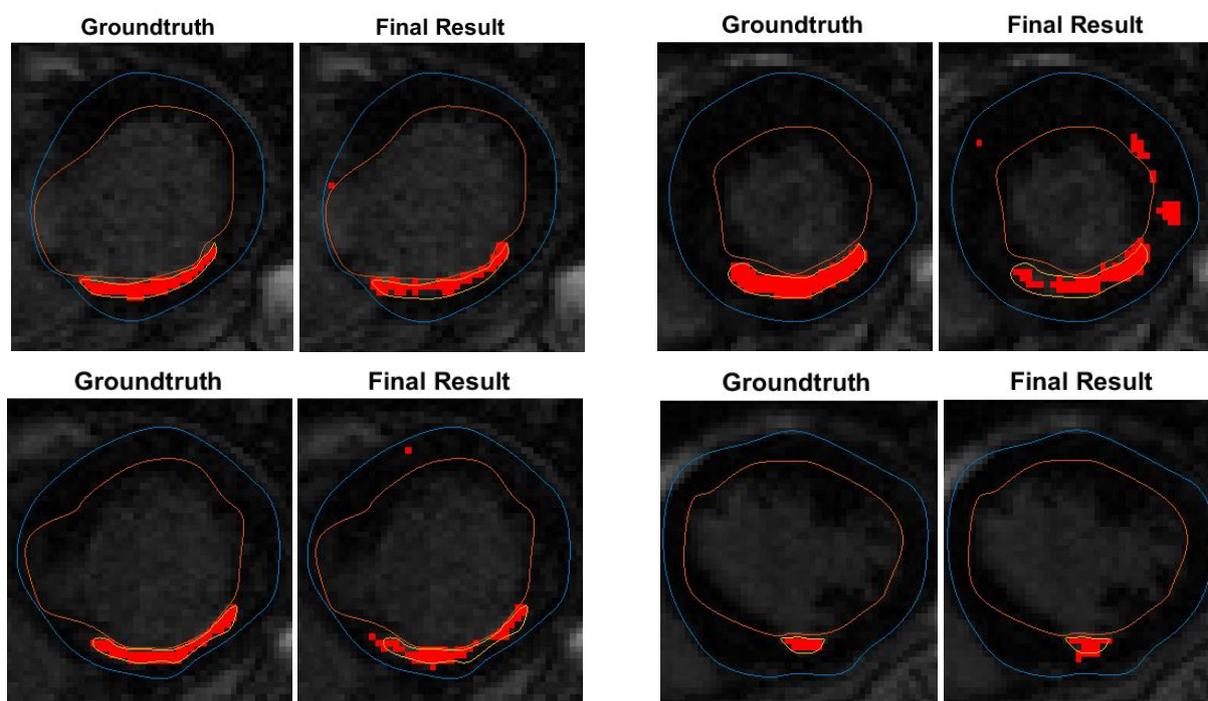
A continuación, se muestran los resultados finales del clasificador agrupados en sus respectivos volúmenes.

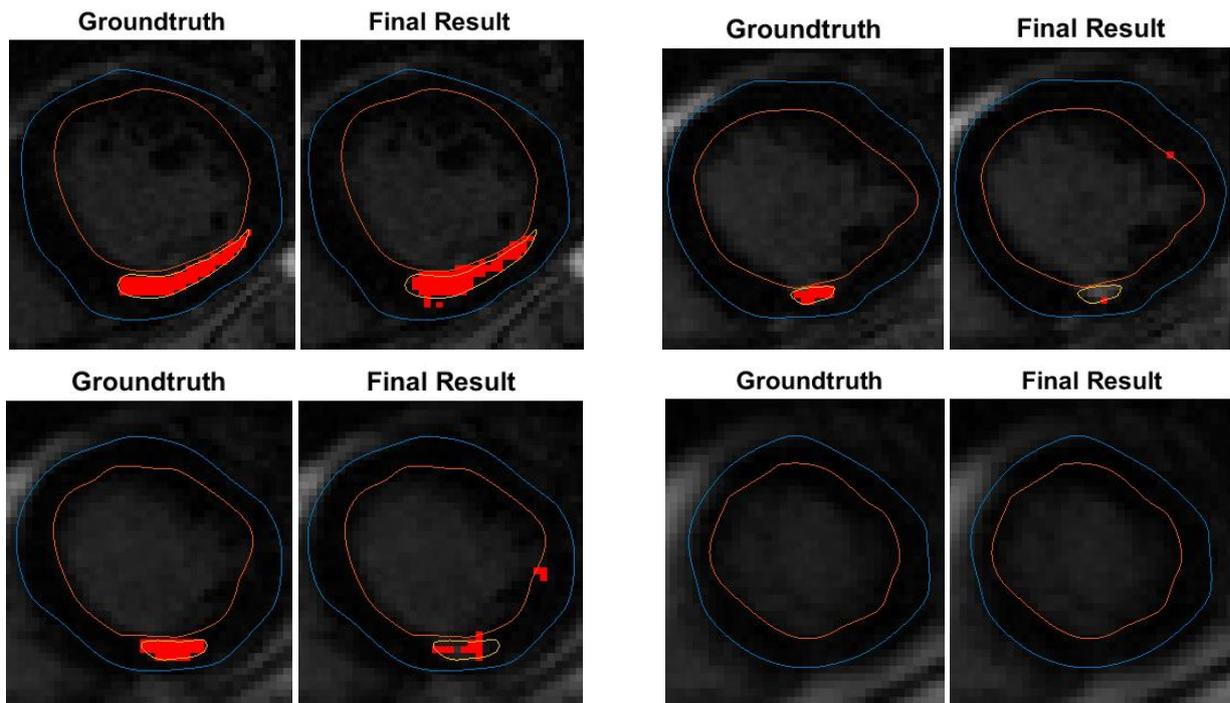
## Volumen 1 (001 – 008)



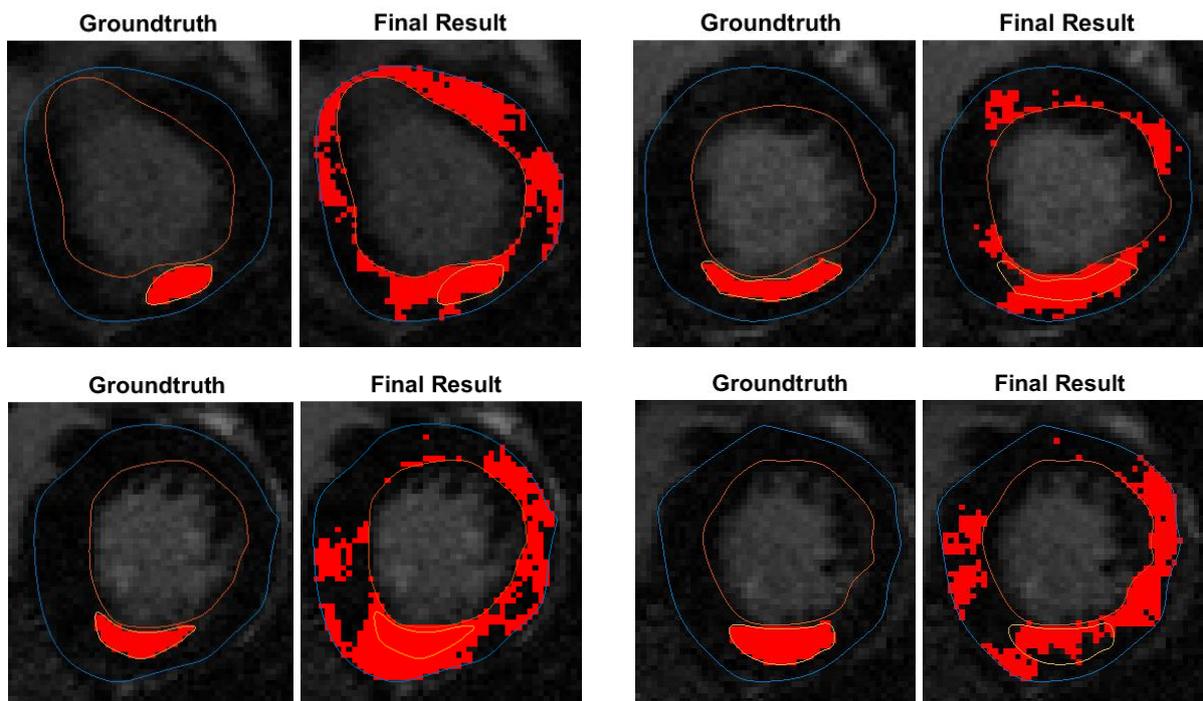


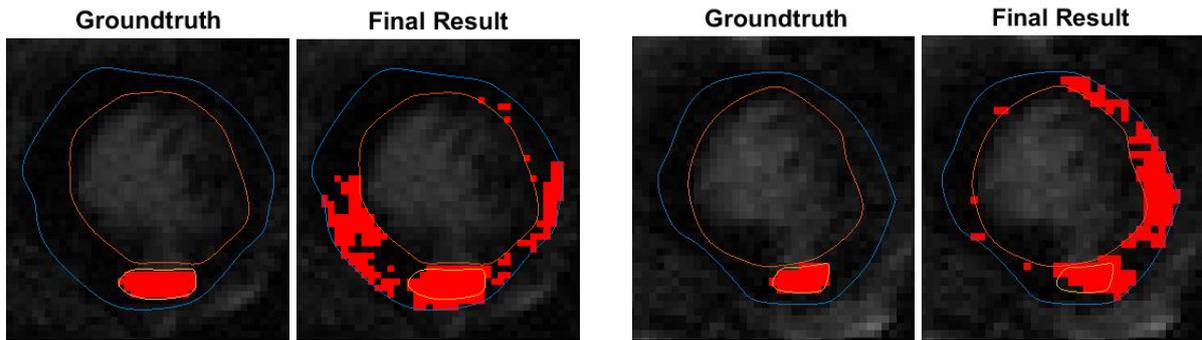
*Volumen 2 (0010 – 017)*



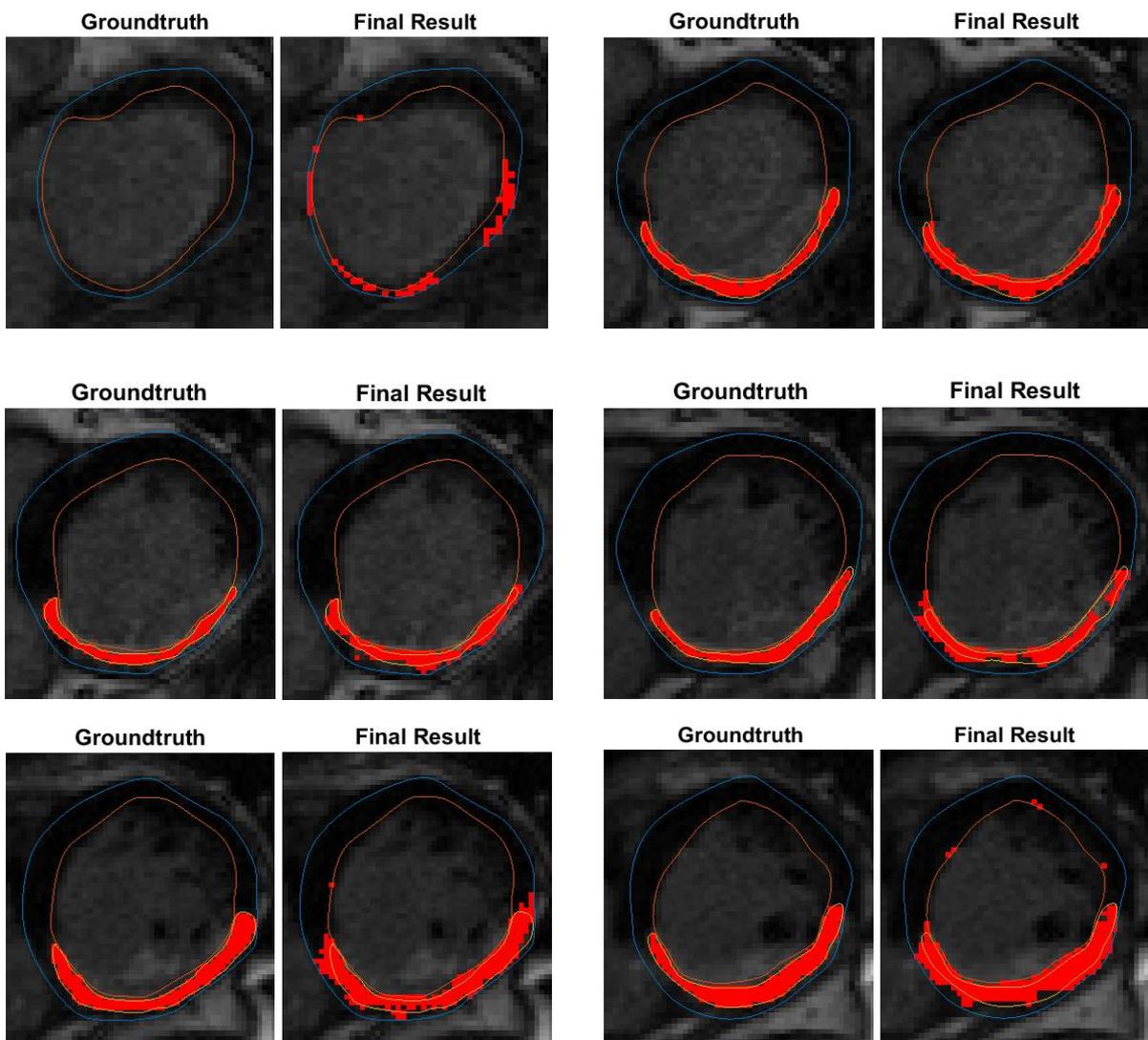


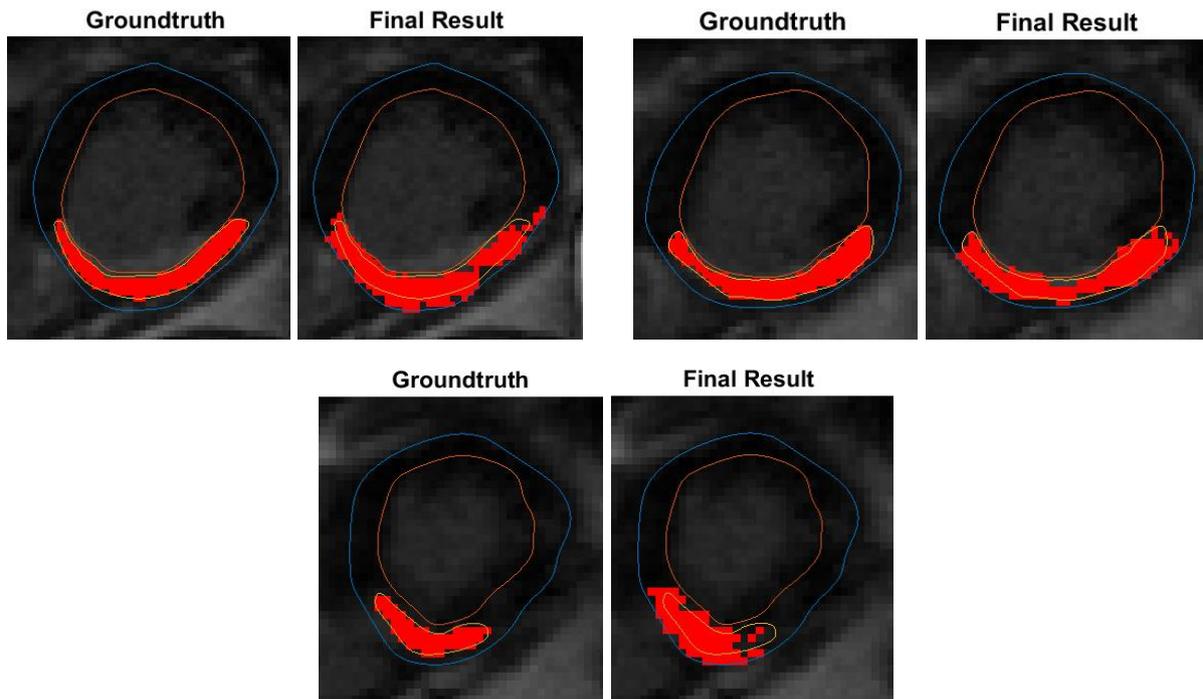
**Volumen 3 (019 – 024)**



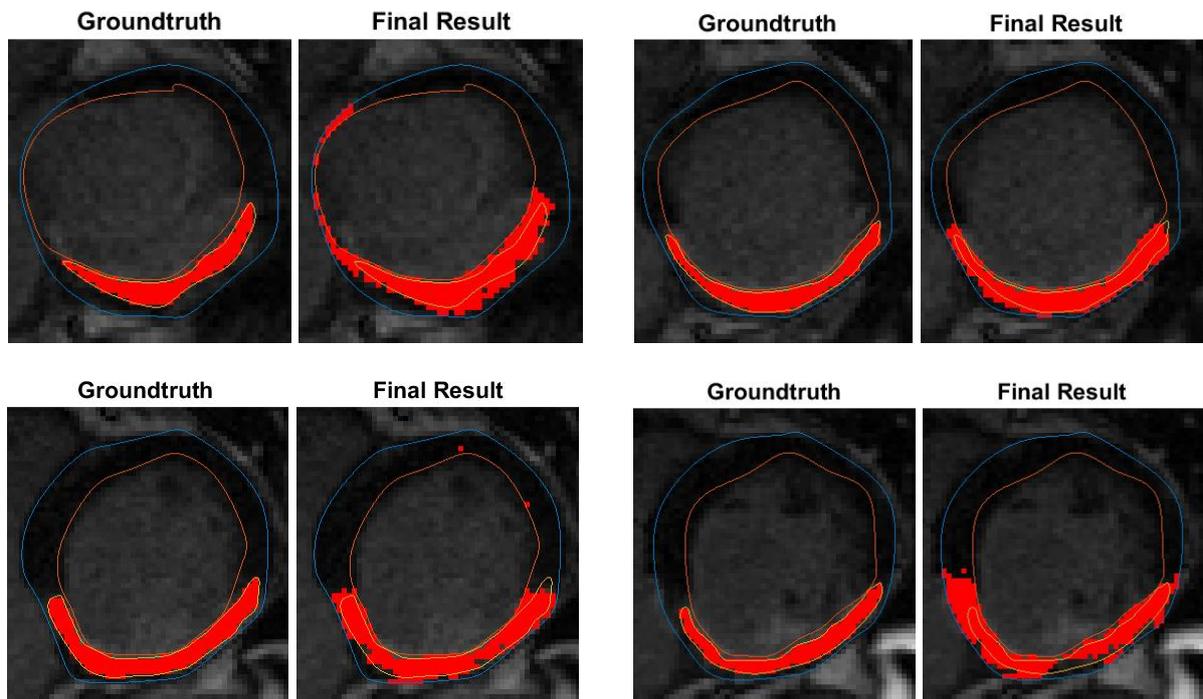


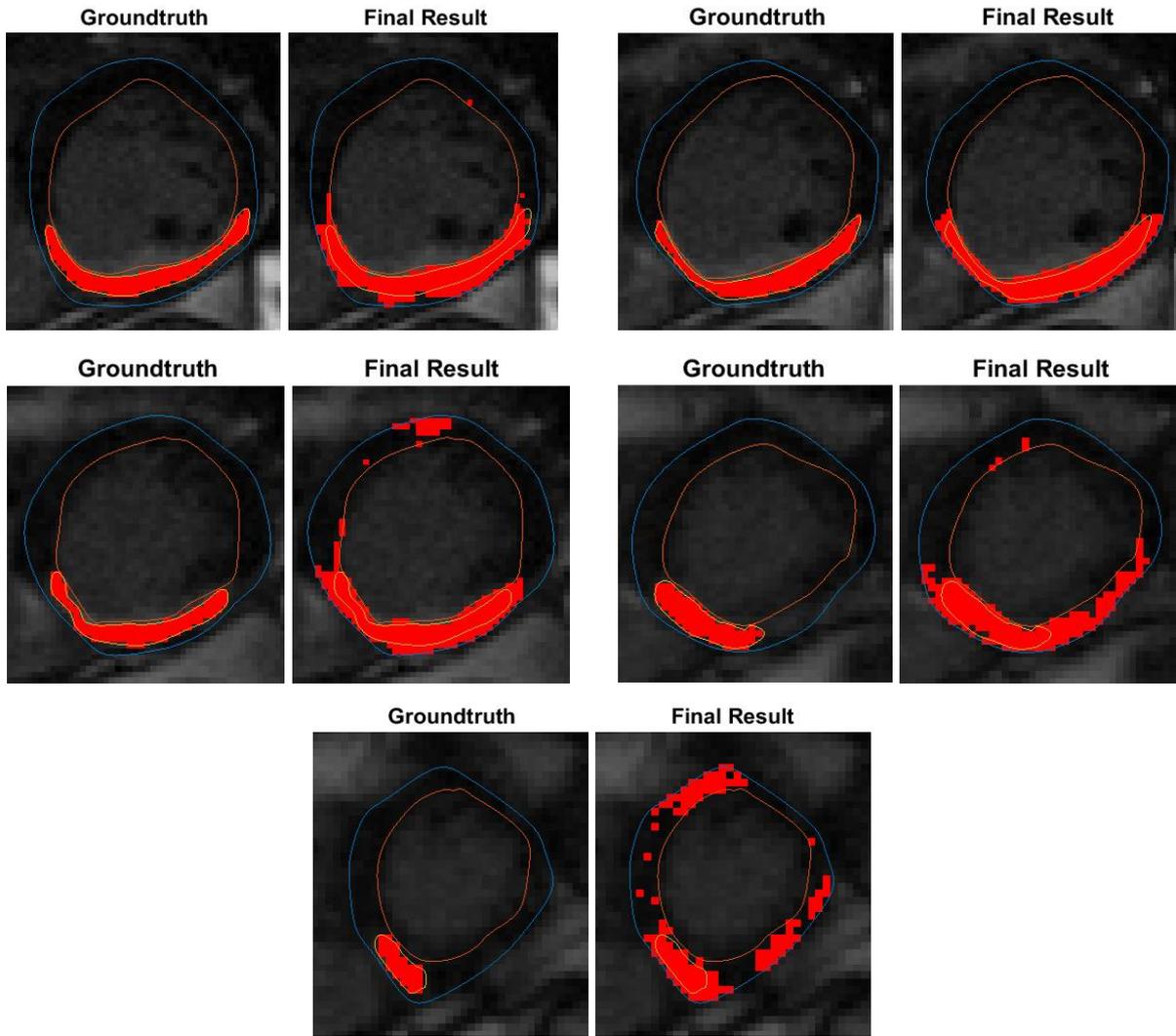
**Volumen 4 (025 – 033)**



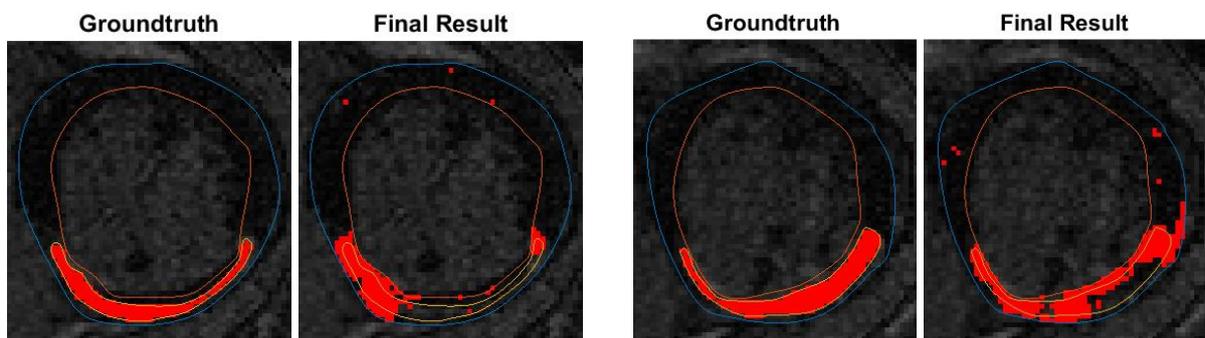


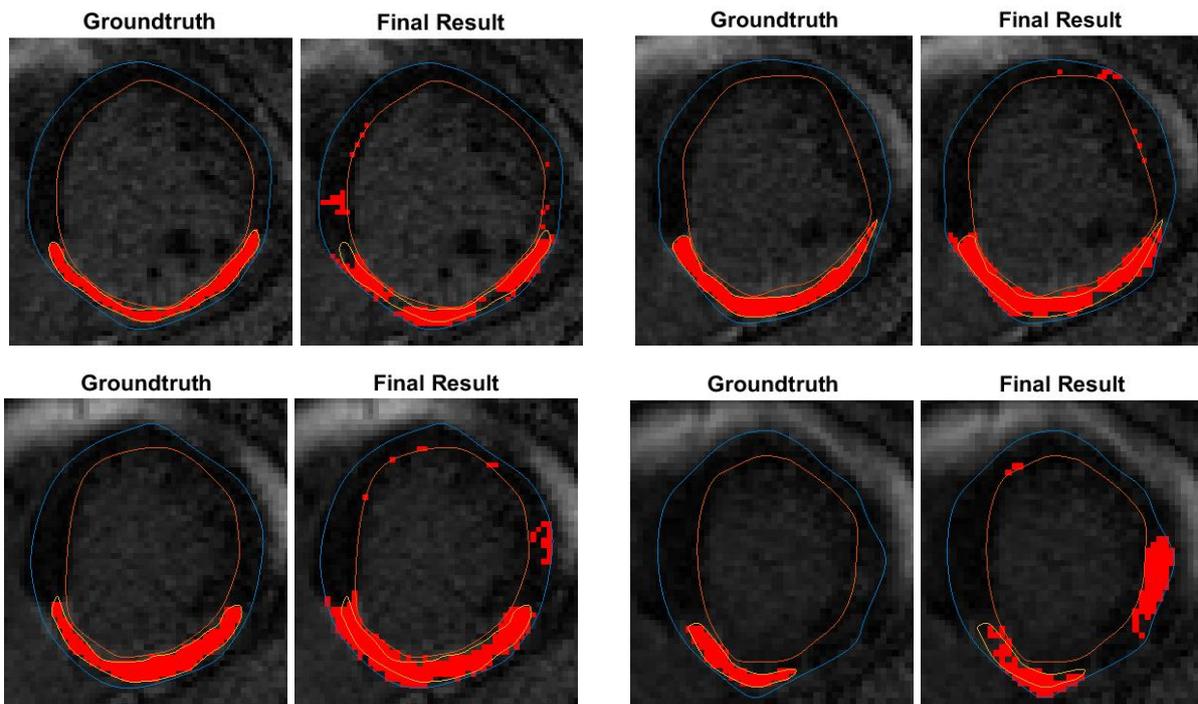
**Volumen 5 (034 – 042)**



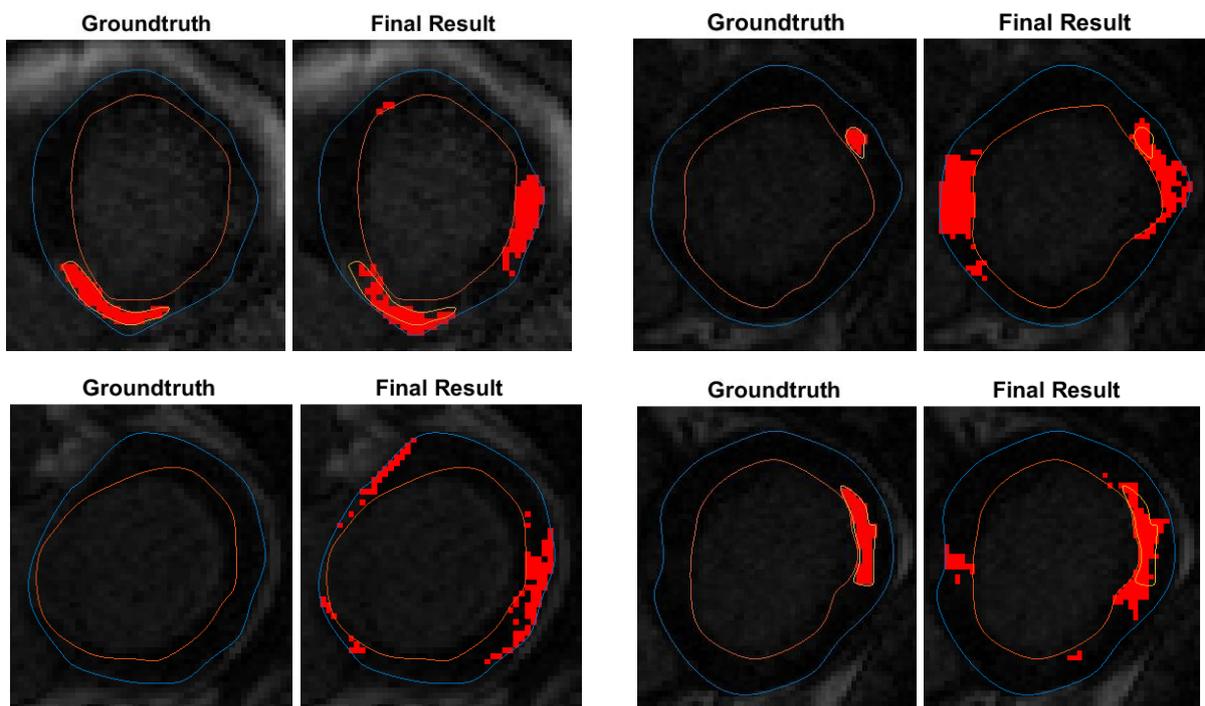


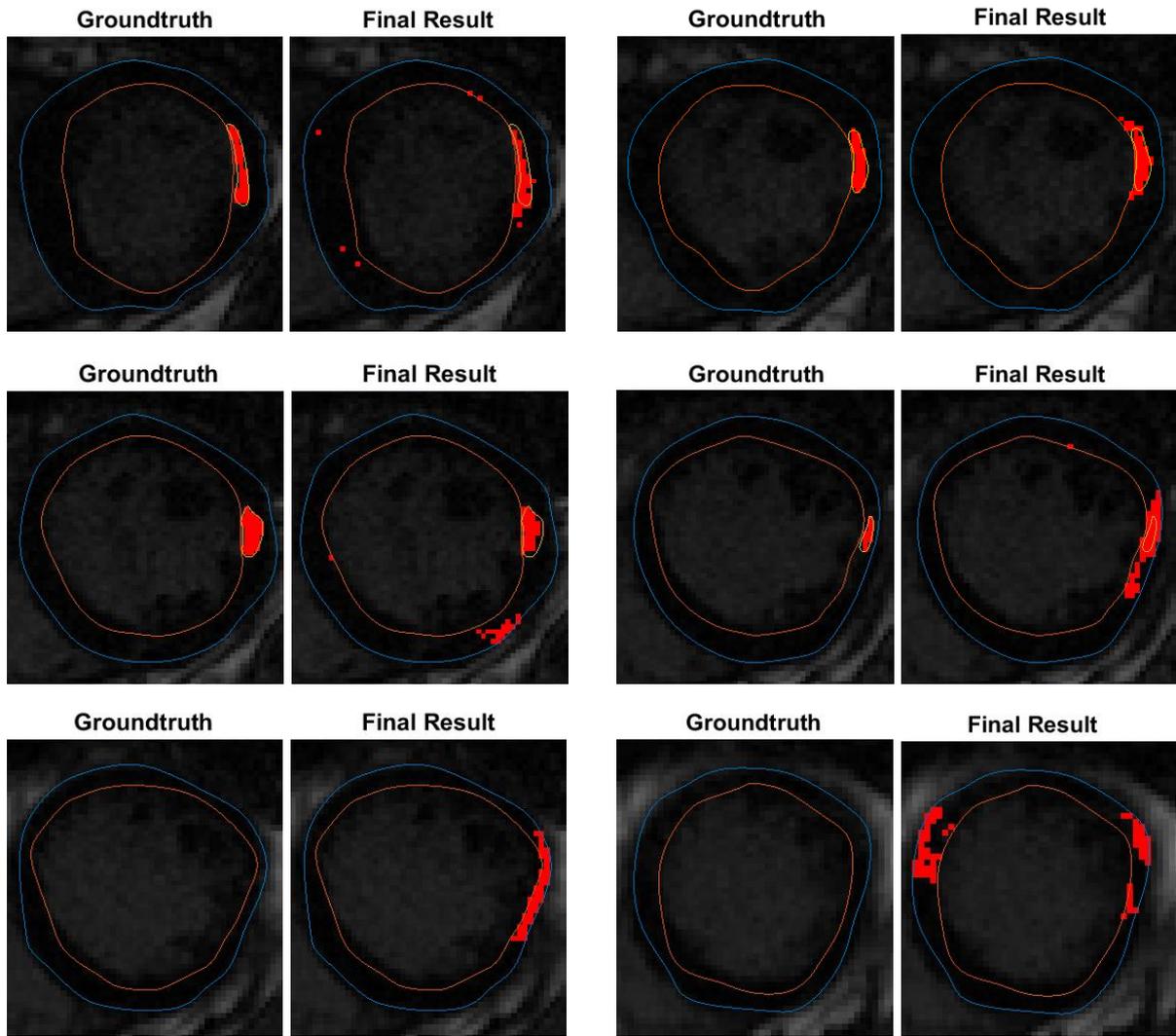
**Volumen 6 (043 – 048)**



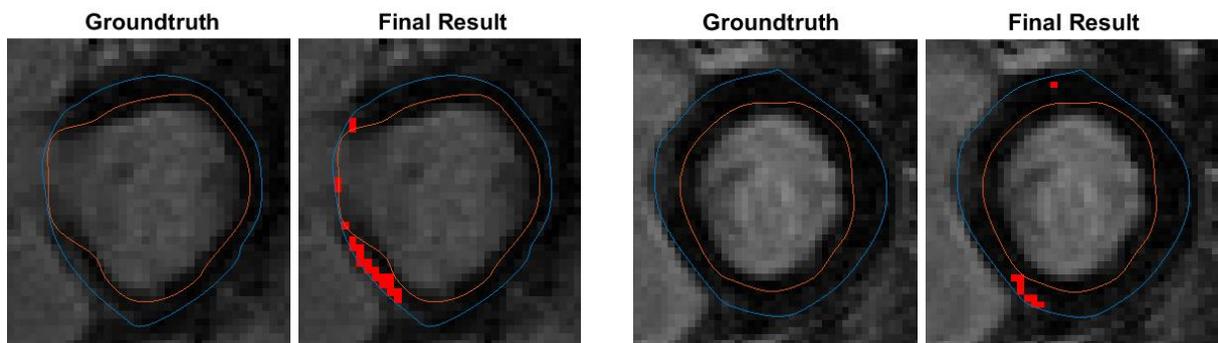


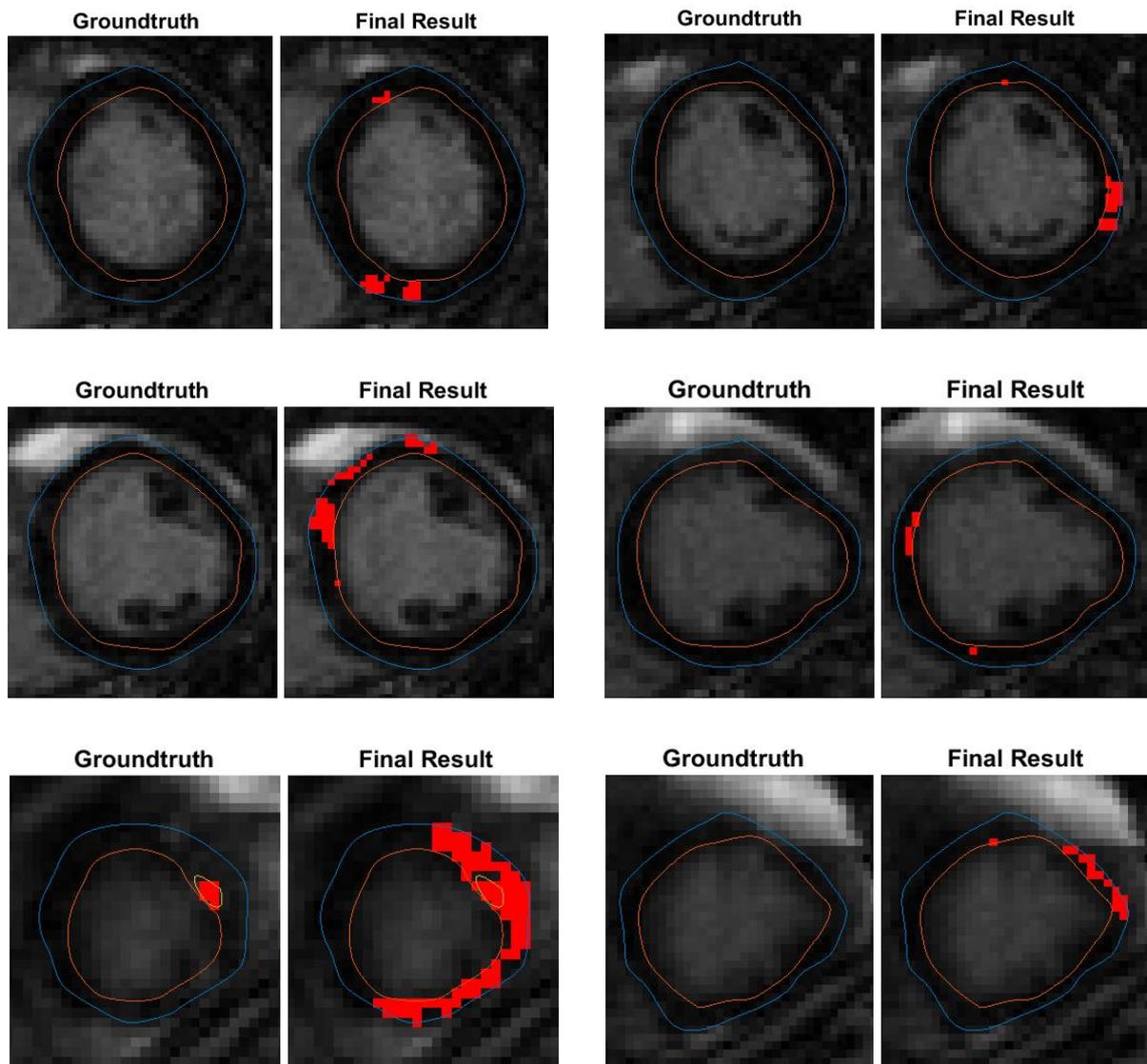
**Volumen 7 (049 – 057)**



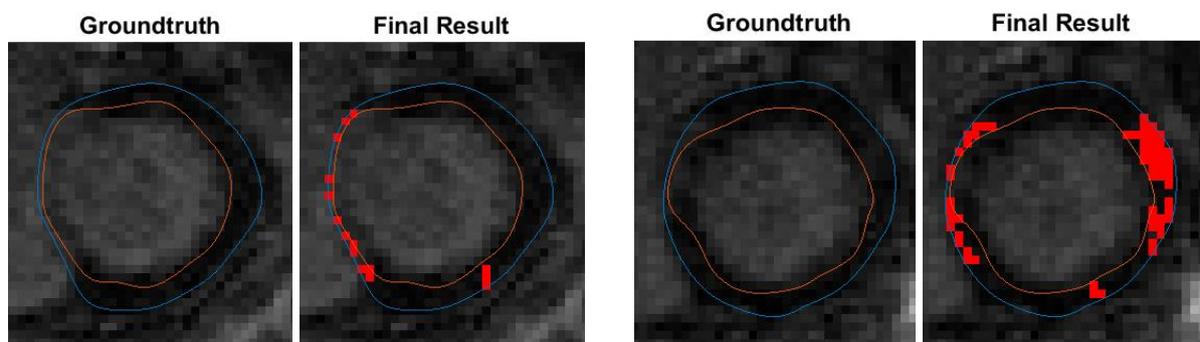


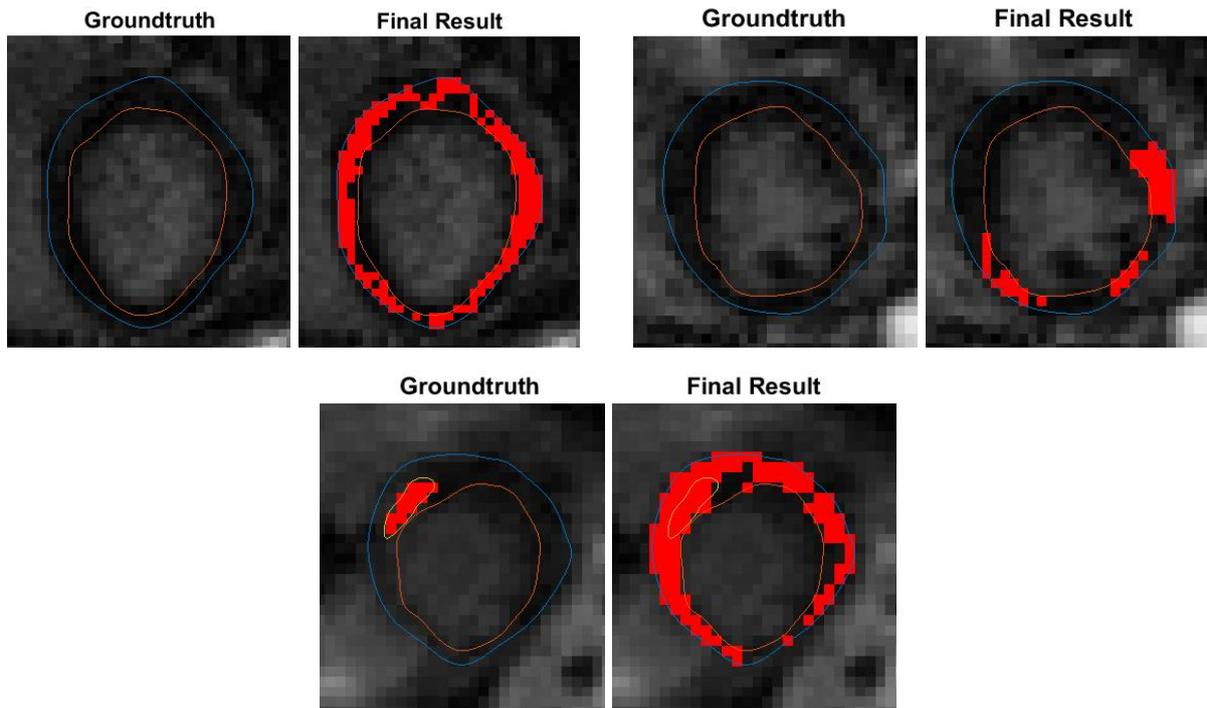
*Volumen 8 (058 – 065)*



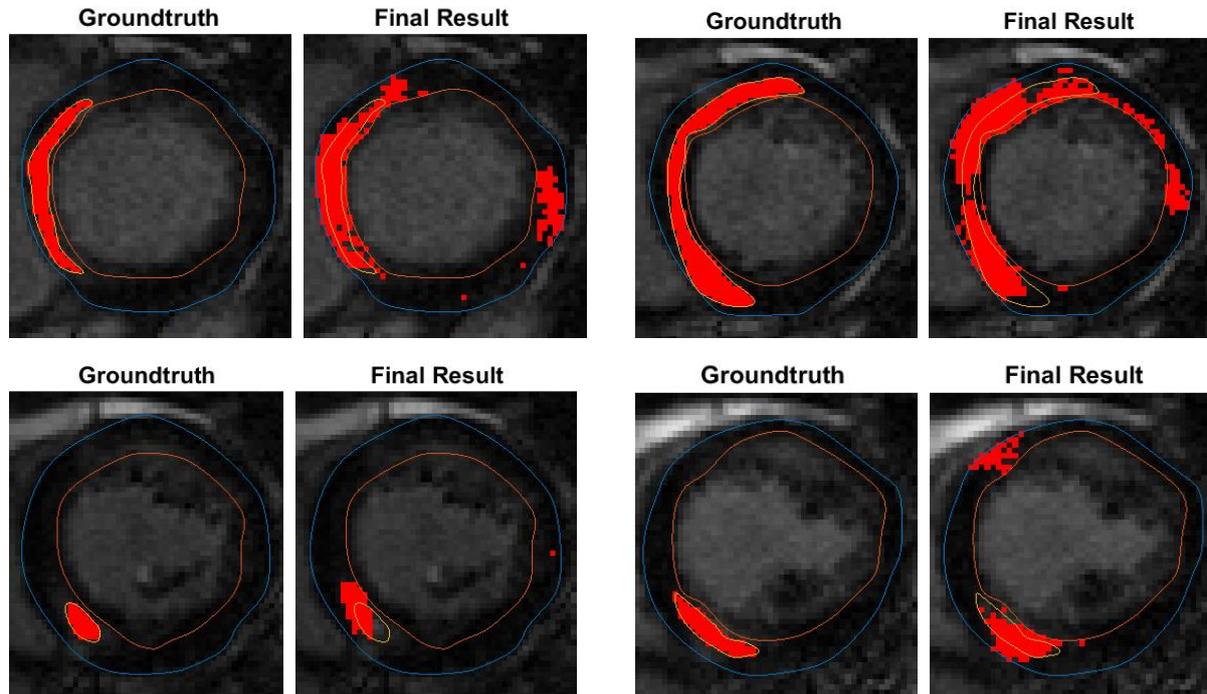


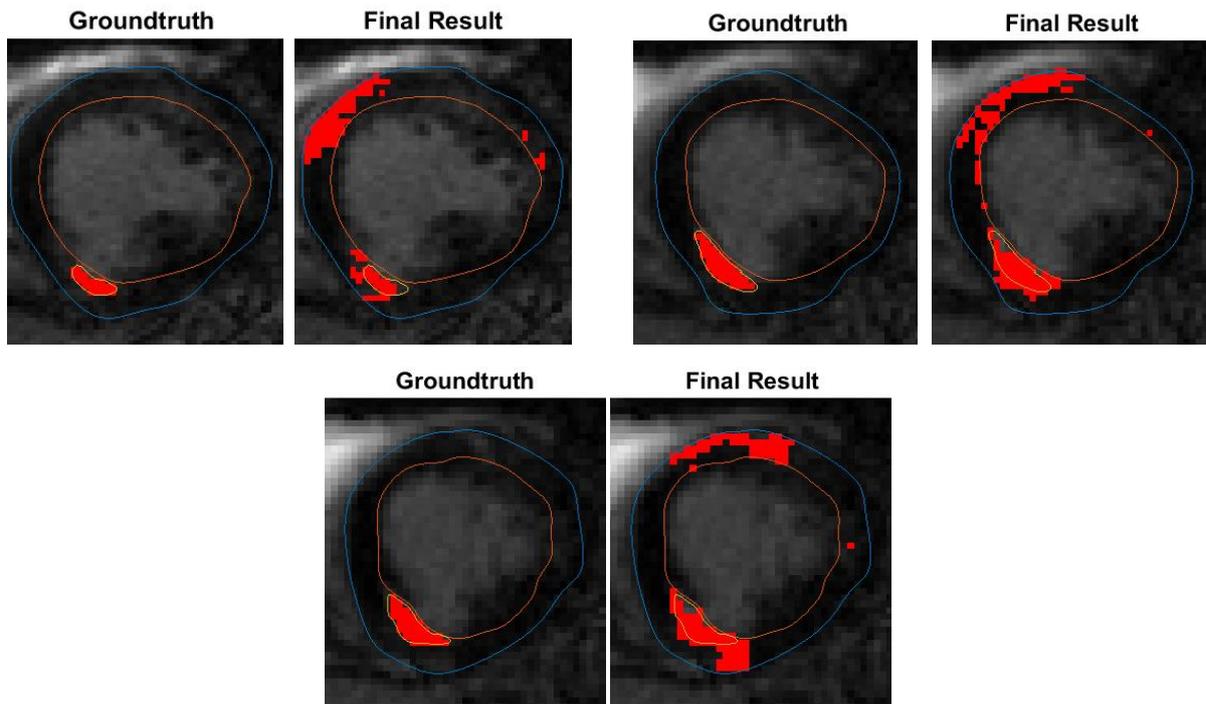
**Volumen 9 (066 – 072)**



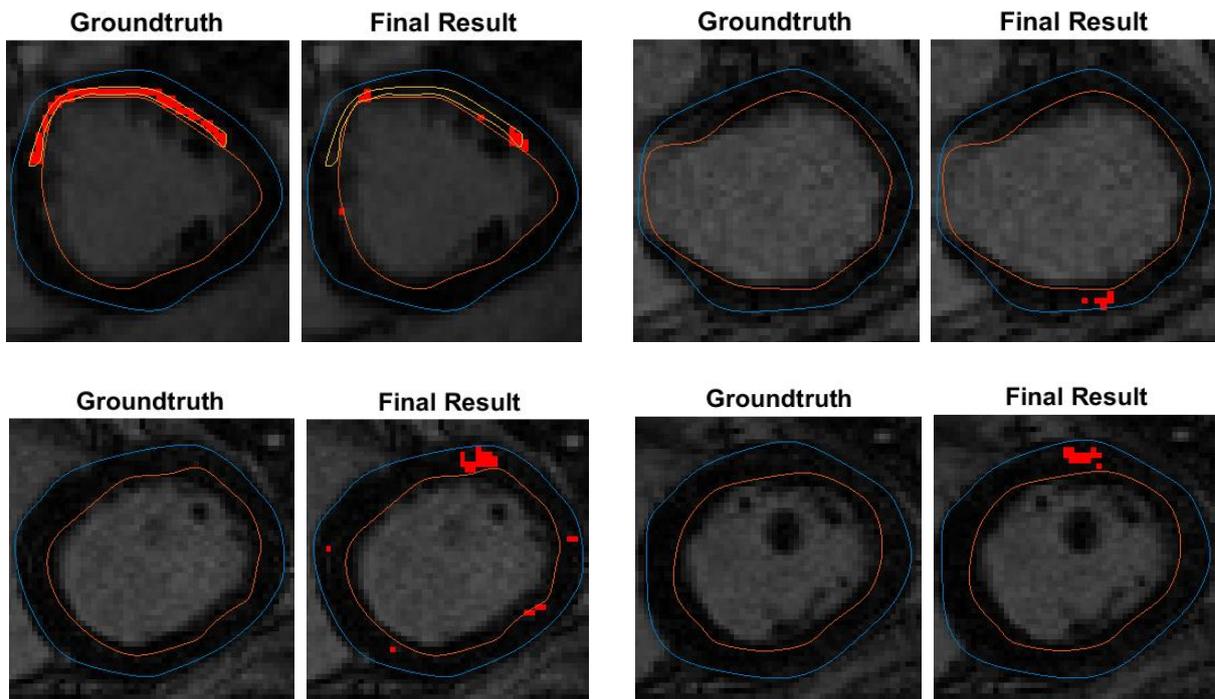


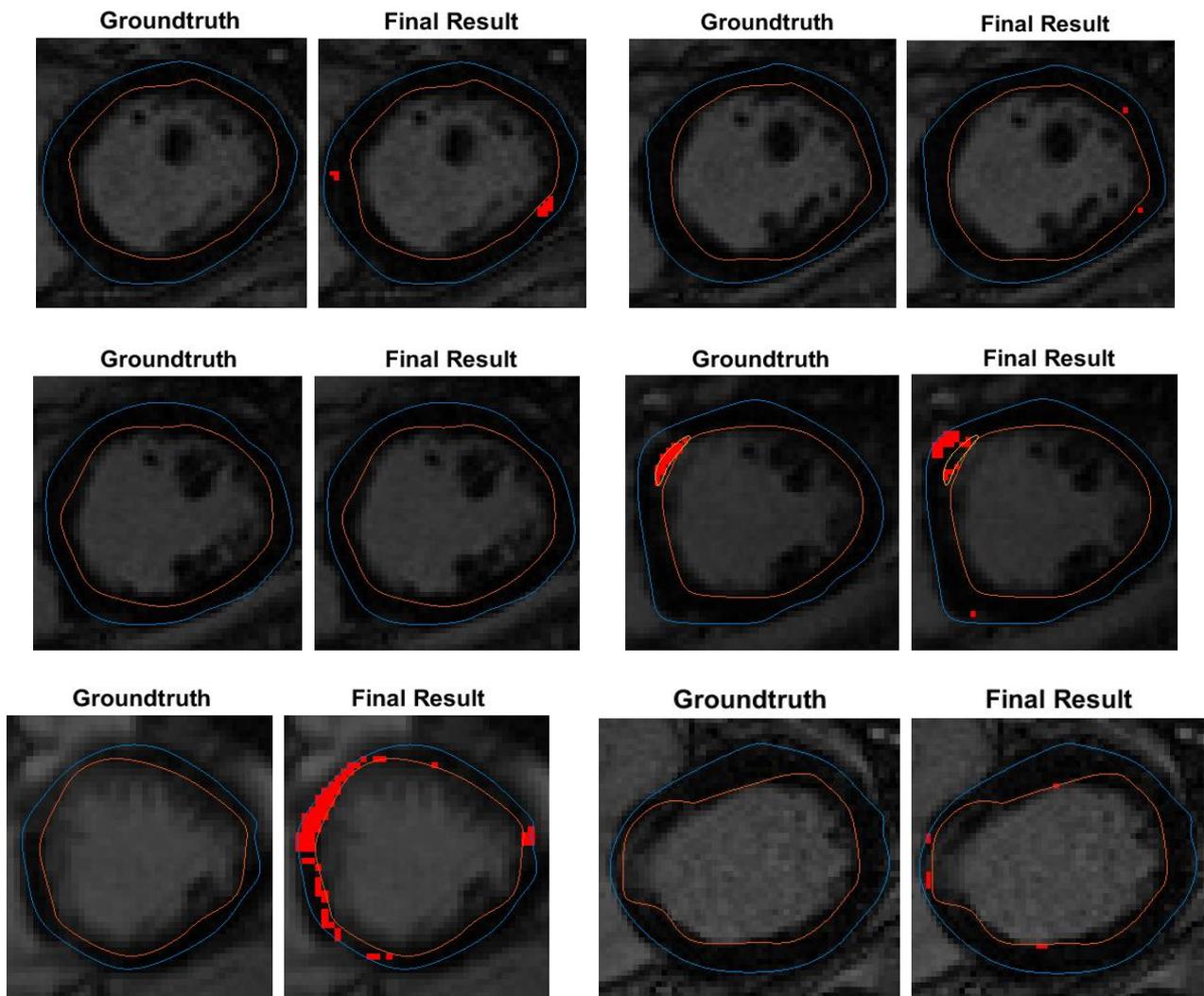
*Volumen 10 (073 – 079)*



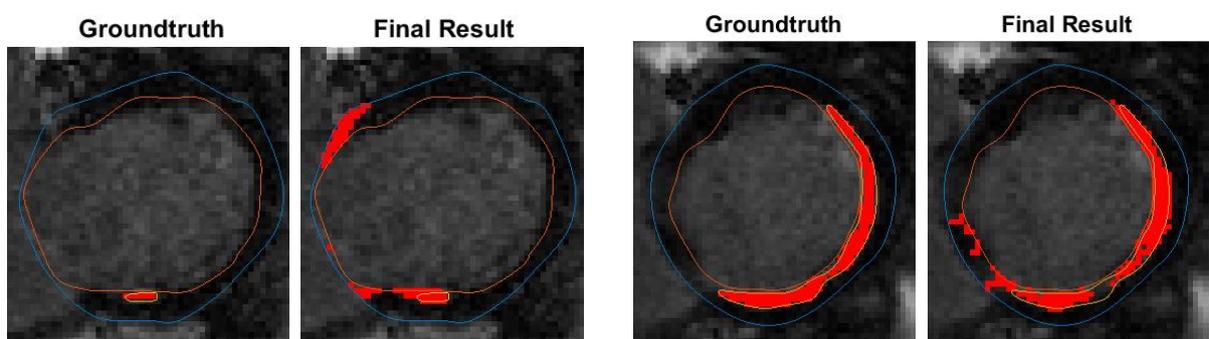


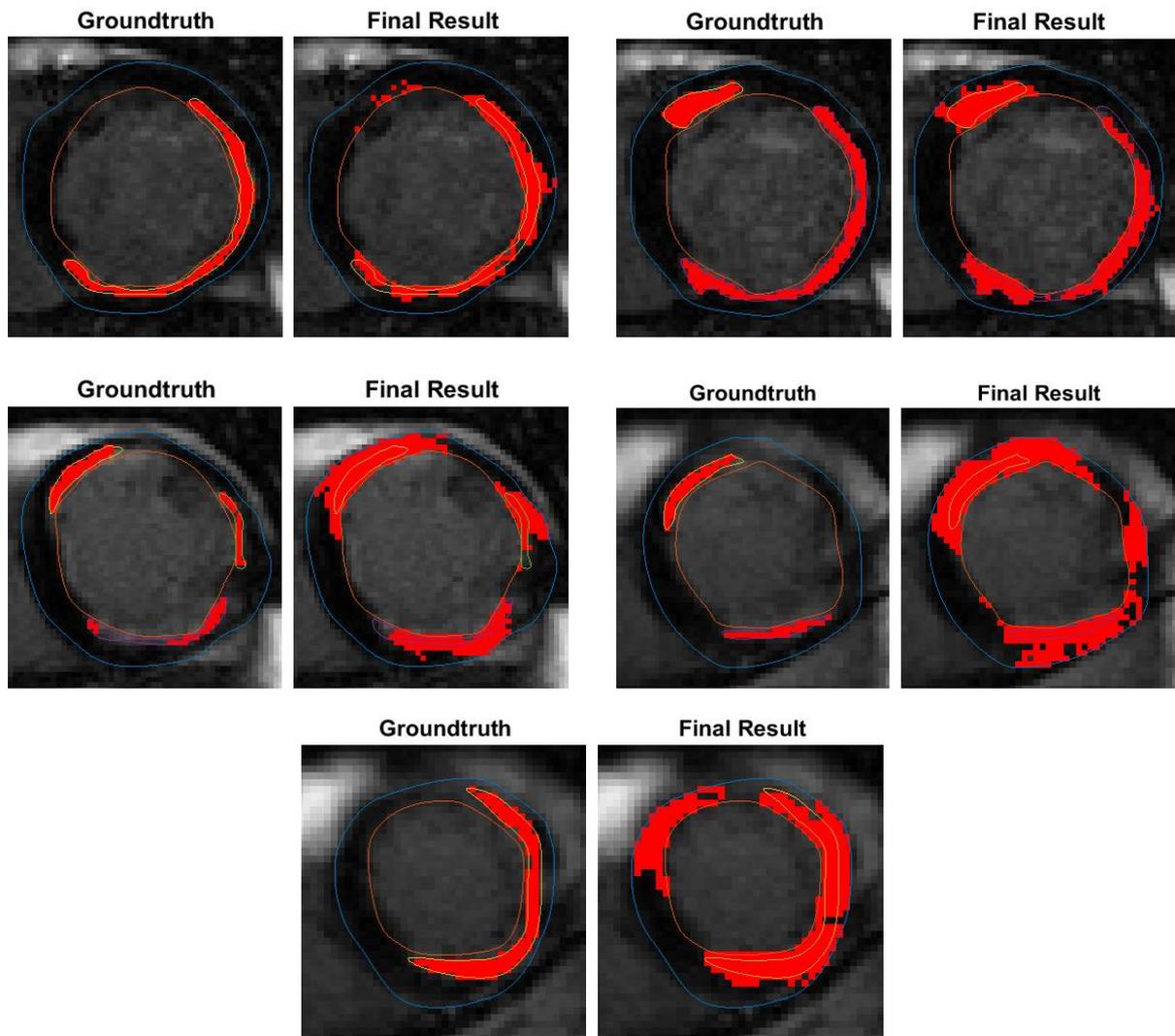
Volumen 11 (081 – 090)



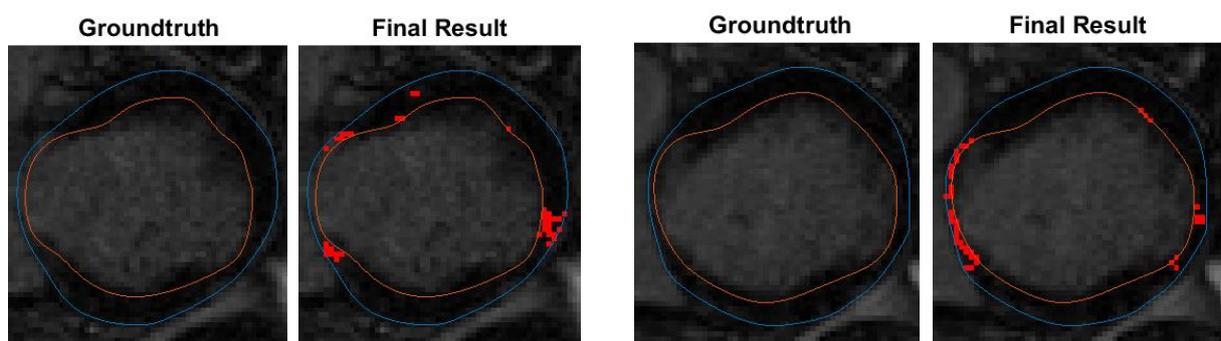


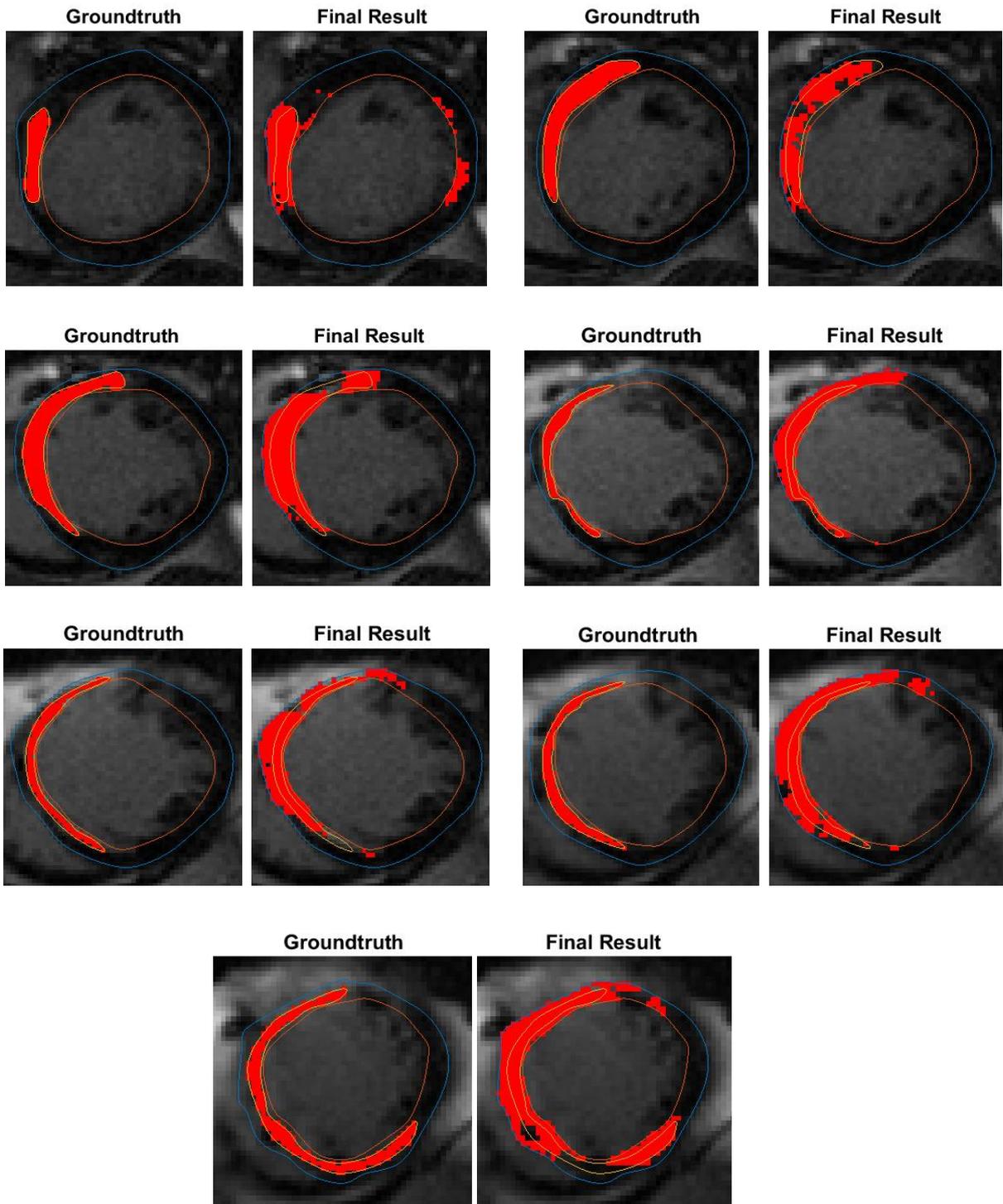
*Volumen 12 (091 – 099)*

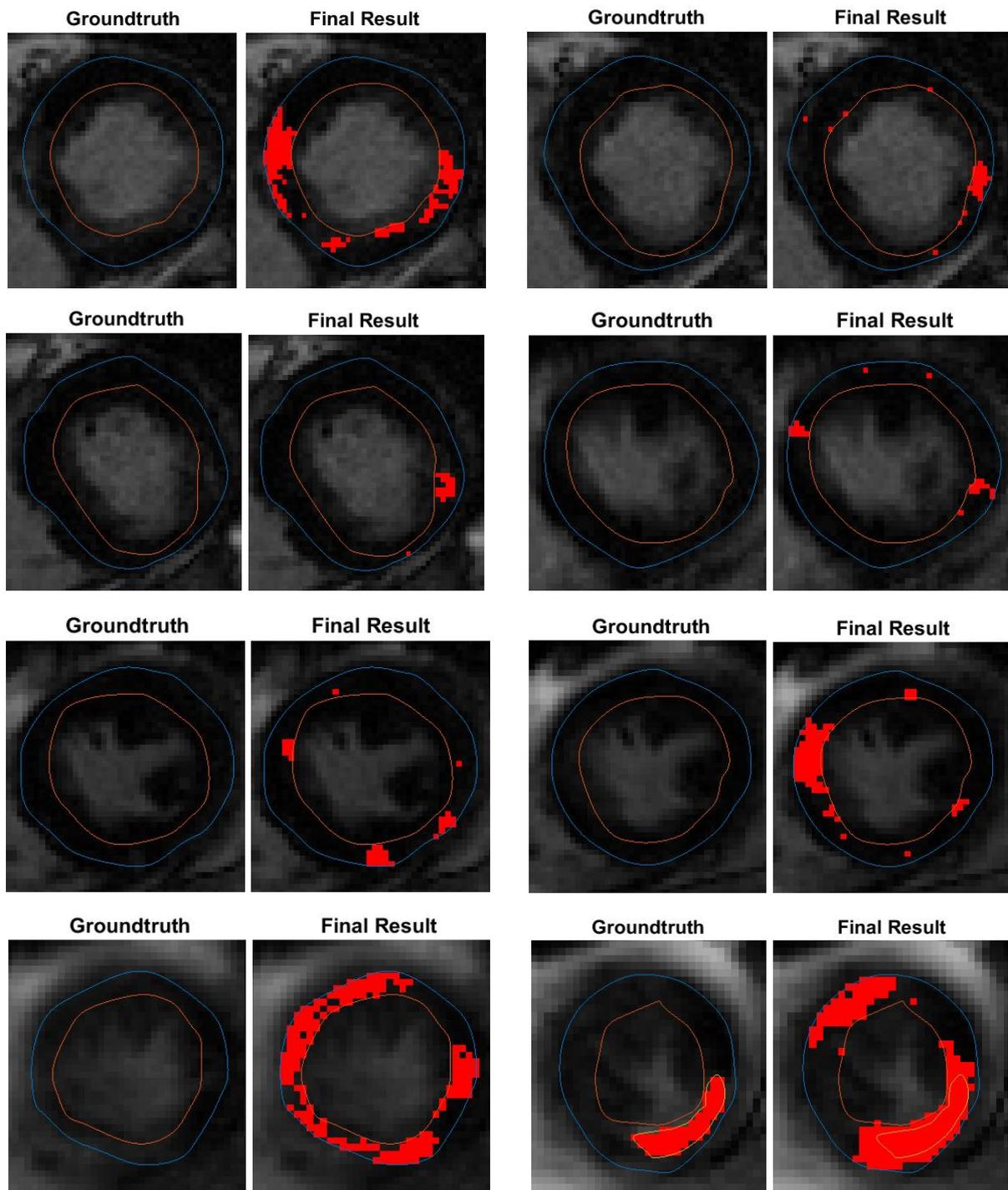




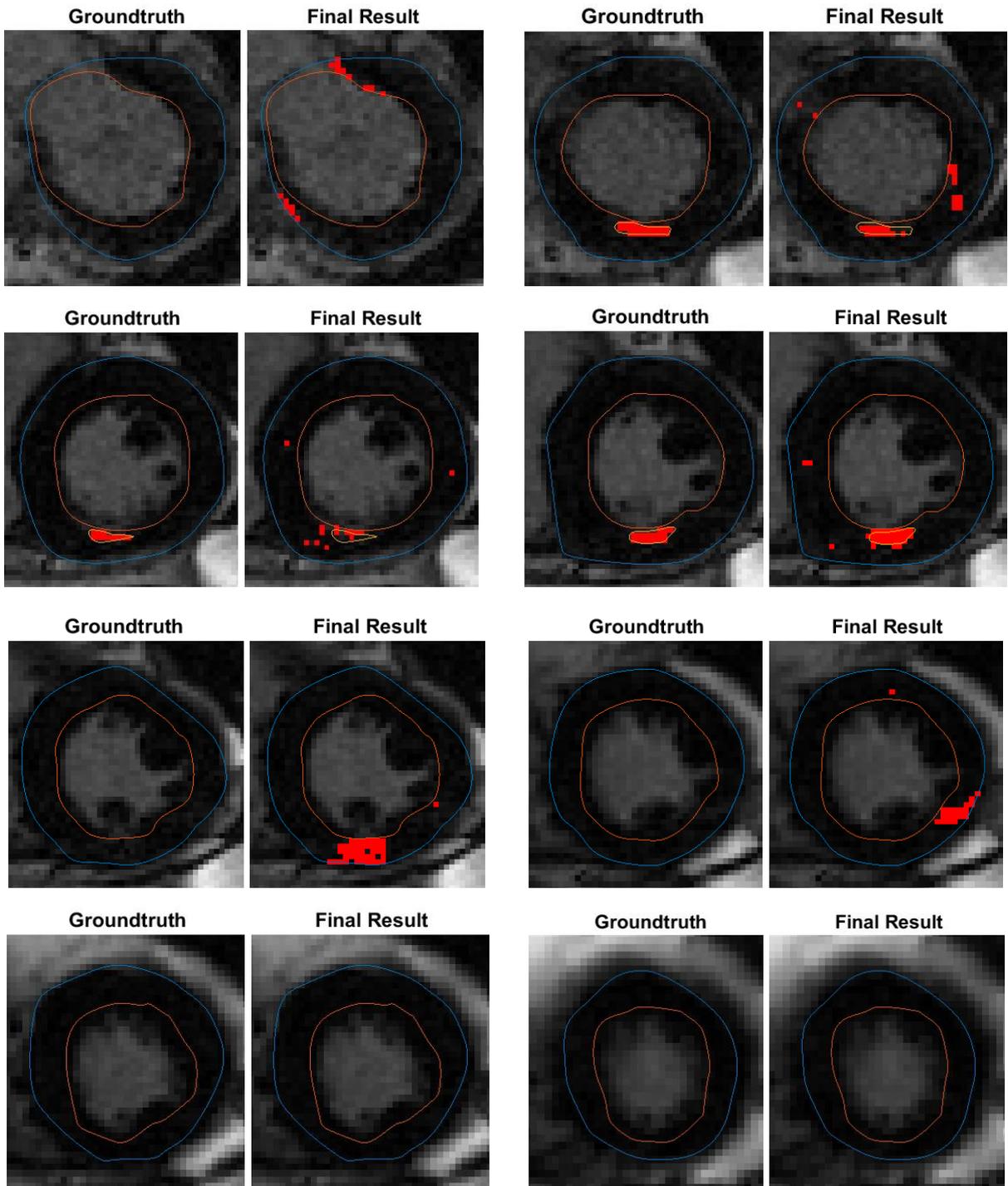
**Volumen 13 (100 – 108)**



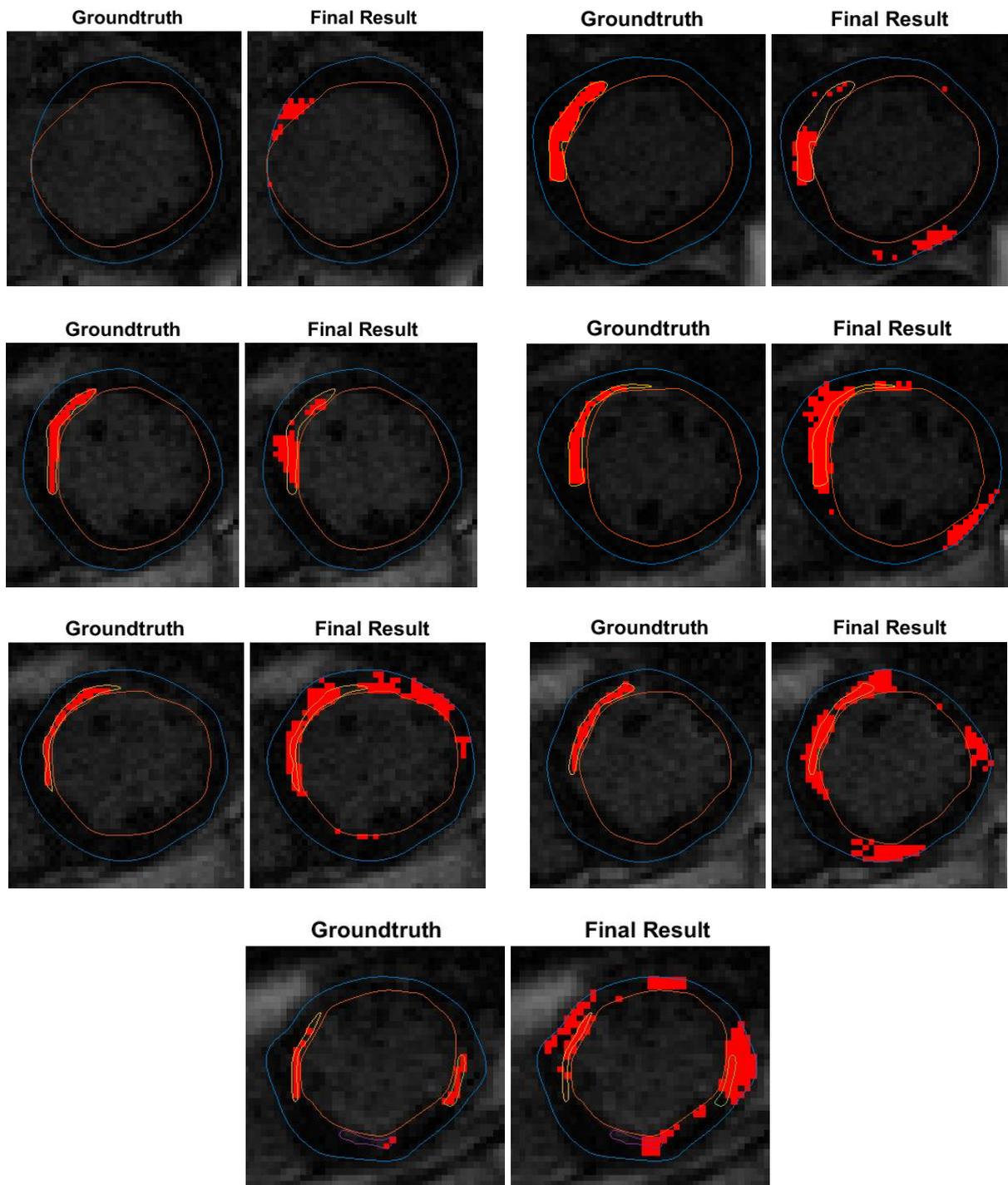


**Volumen 14 (109 – 116)**

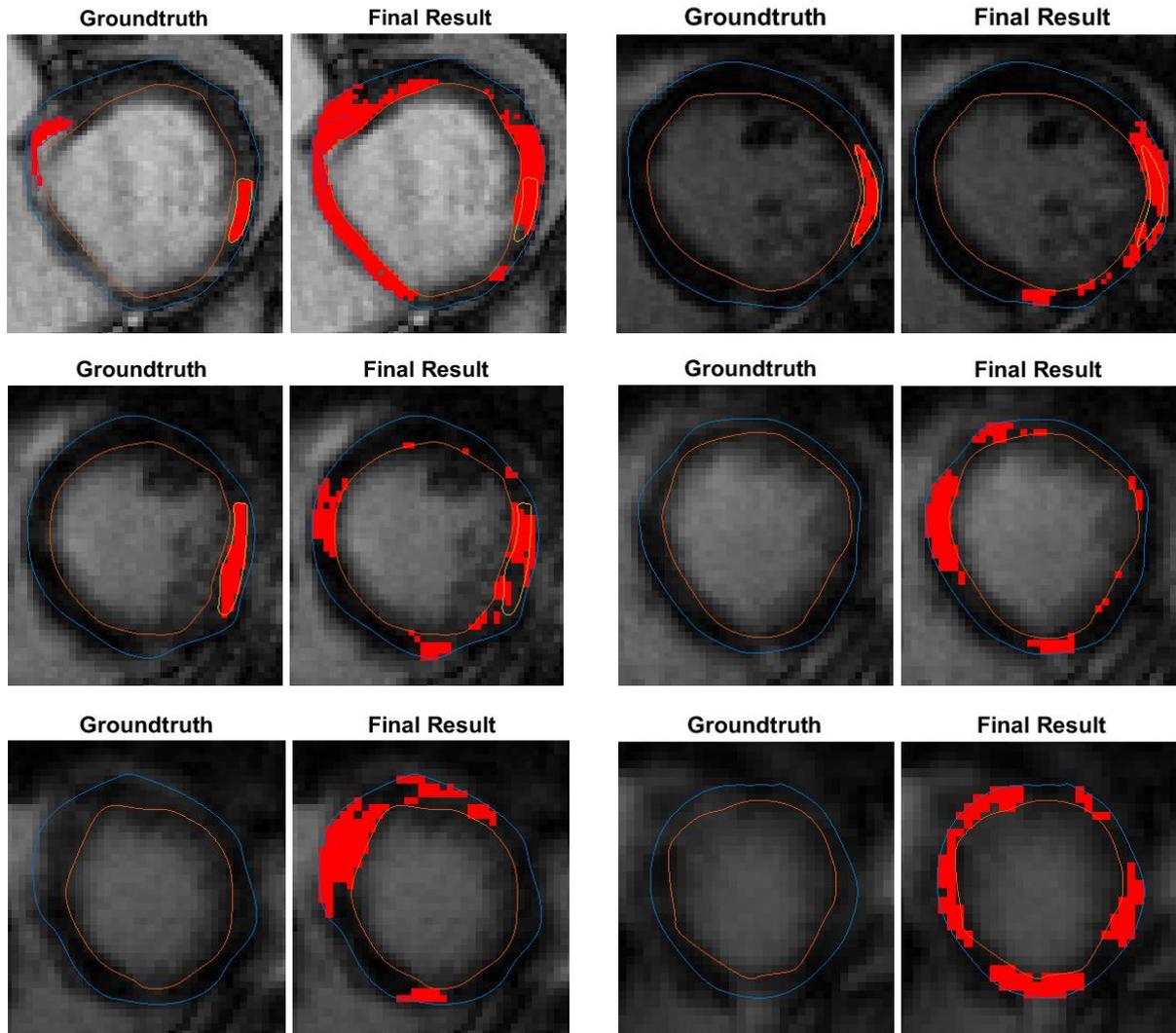
**Volumen 15 (117 – 124)**



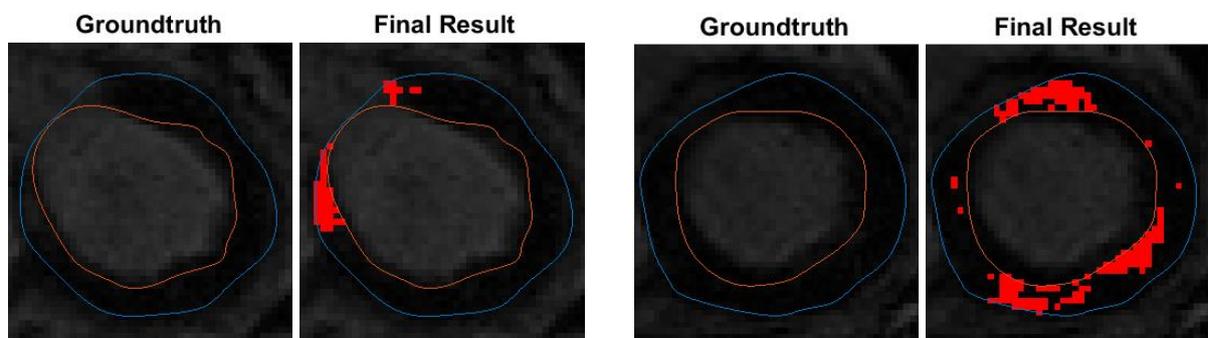
**Volumen 16 (125 – 132)**

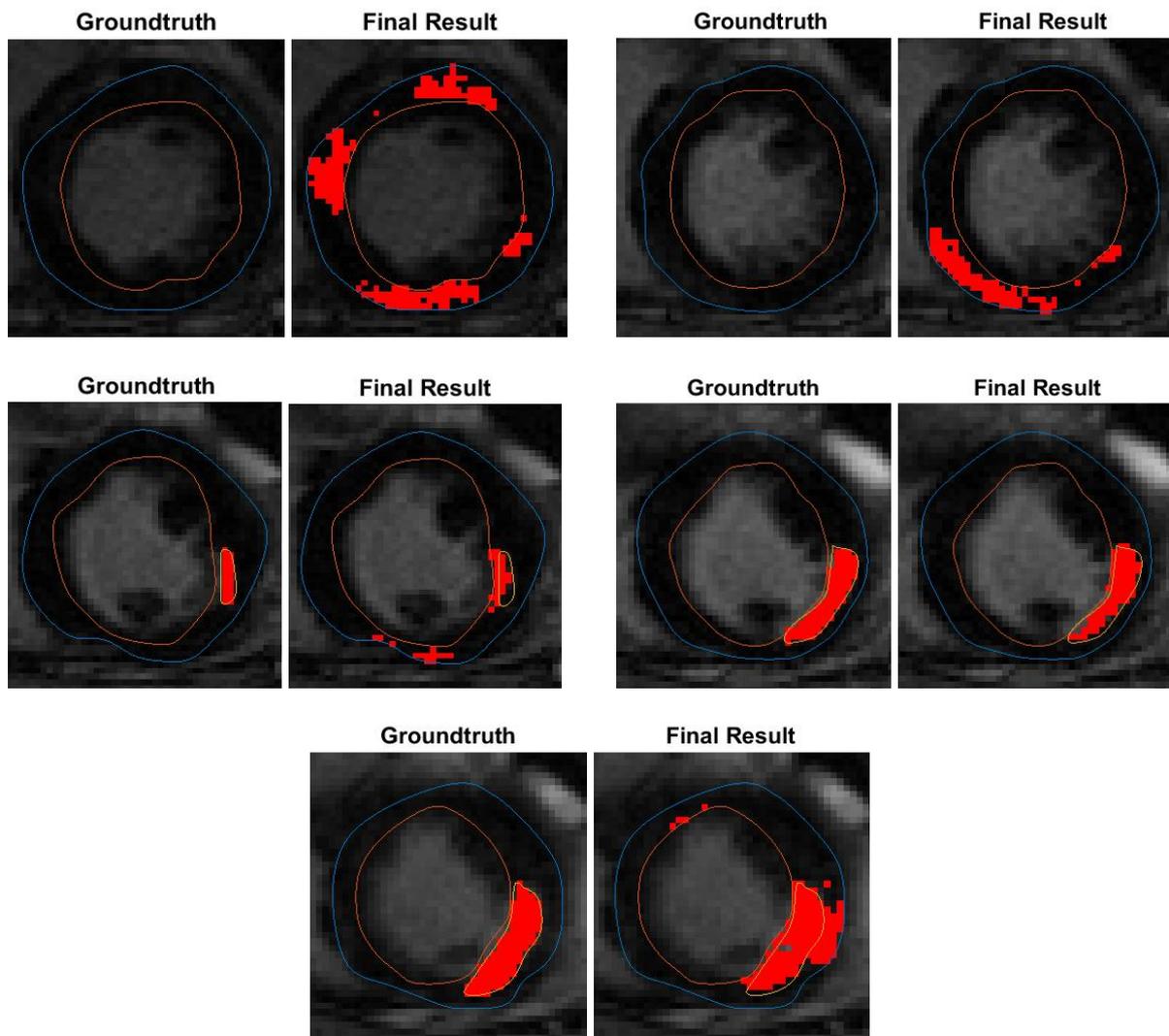


**Volumen 17 (133 – 140)**

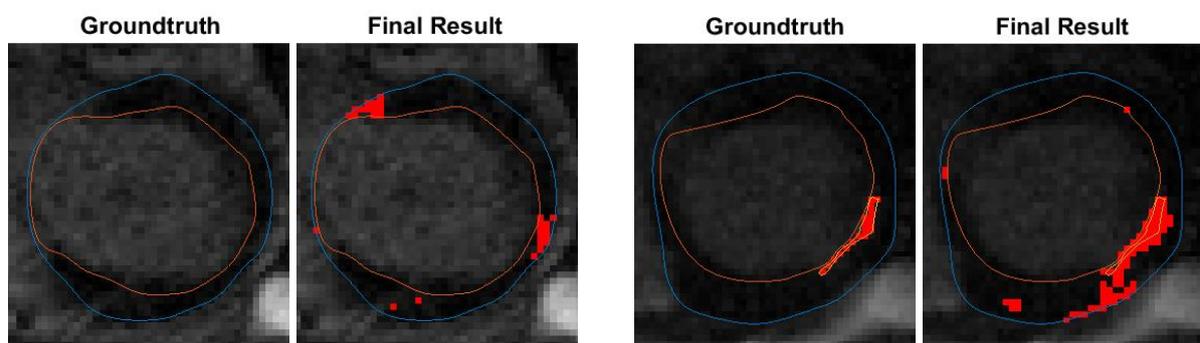


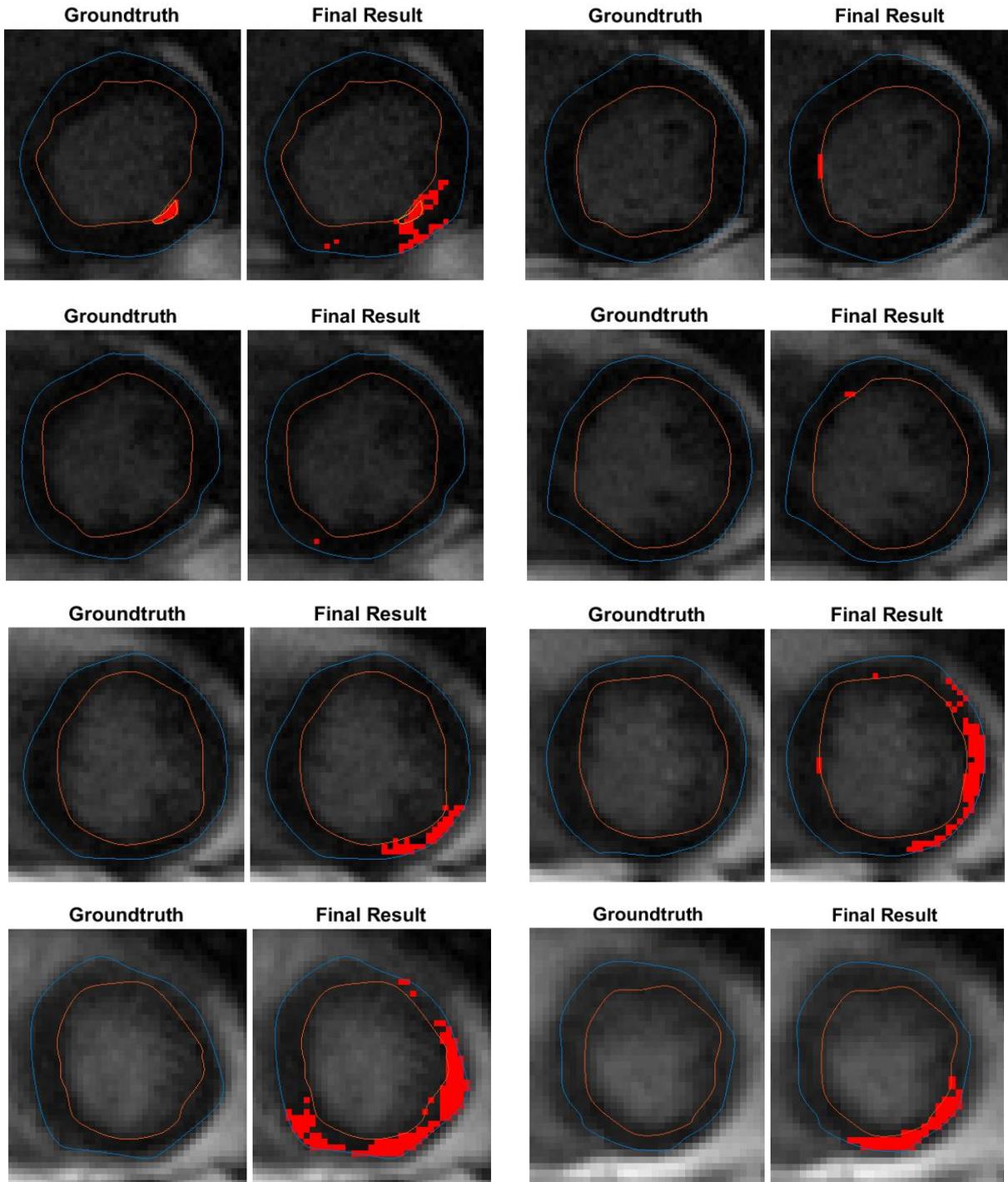
**Volumen 18 (141 – 147)**





**Volumen 19 (148 – 157)**





**Volumen 20 (158 – 164)**

