



# Segmentación de Estructuras Cardiovasculares mediante Redes Neuronales Convolucionales



Tutora: Dra. Jimena Olveres Montiel

Steve Alejandro Avendaño García

A large, thick, black L-shaped frame that starts at the top-left, goes right, then down, then right again, forming a partial rectangular border around the central text.

# JUSTIFICACIÓN

- Las enfermedades cardiovasculares representan la primera causa de muerte a nivel mundial y en occidente.

Principales causas de muerte, 2016



Fig. 1 Estimaciones de salud global 2016. Organización Mundial de la Salud 2018

- La segmentación de todo el volumen cardiaco se realiza con la finalidad de analizar lesiones coronarias o algún tipo de malformación cardiaca, además de que es un pre-requisito para crear modelos anatómicos que permitan llevar a cabo la planeación de alguna intervención quirúrgica.

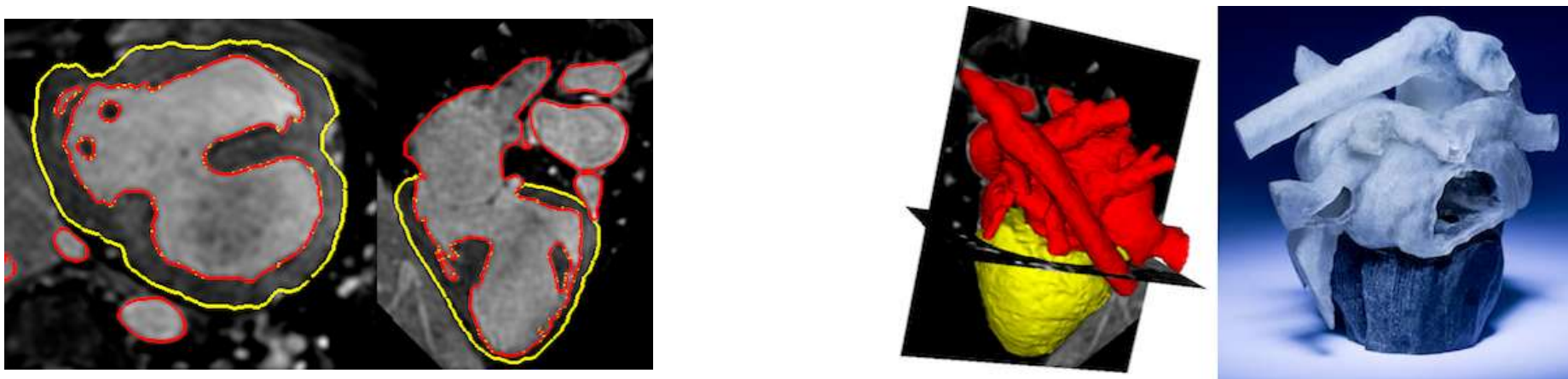


Fig. 2 Segmentación y reconstrucción del volumen cardiaco.

- La segmentación manual de las estructuras cardiacas es el método más común, sin embargo, realizar esto en muchas imágenes resulta una tarea laboriosa que demanda mucho tiempo de un médico especialista.

- Estudio:
  - Imagen por RM T1
  - 9 Observadores independientes
  - Tumor Cerebral
  
- La implementación de modelos semiautomáticos y automáticos aun presenta un problema desafiante debido a características como la complejidad y variabilidad en la morfología.

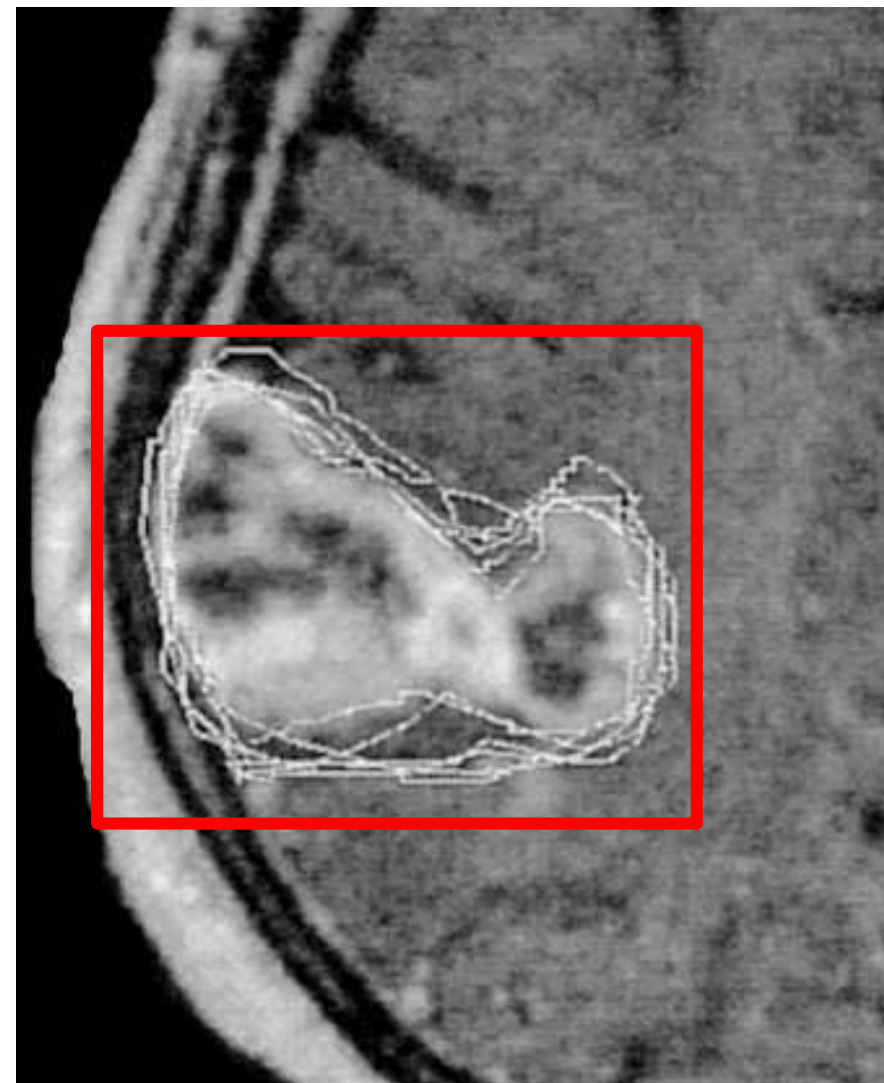


Fig. 3 Variabilidad en la segmentación de imágenes. L. Joskowicz 2019.

## Asistencia por computadora en la imagenología médica

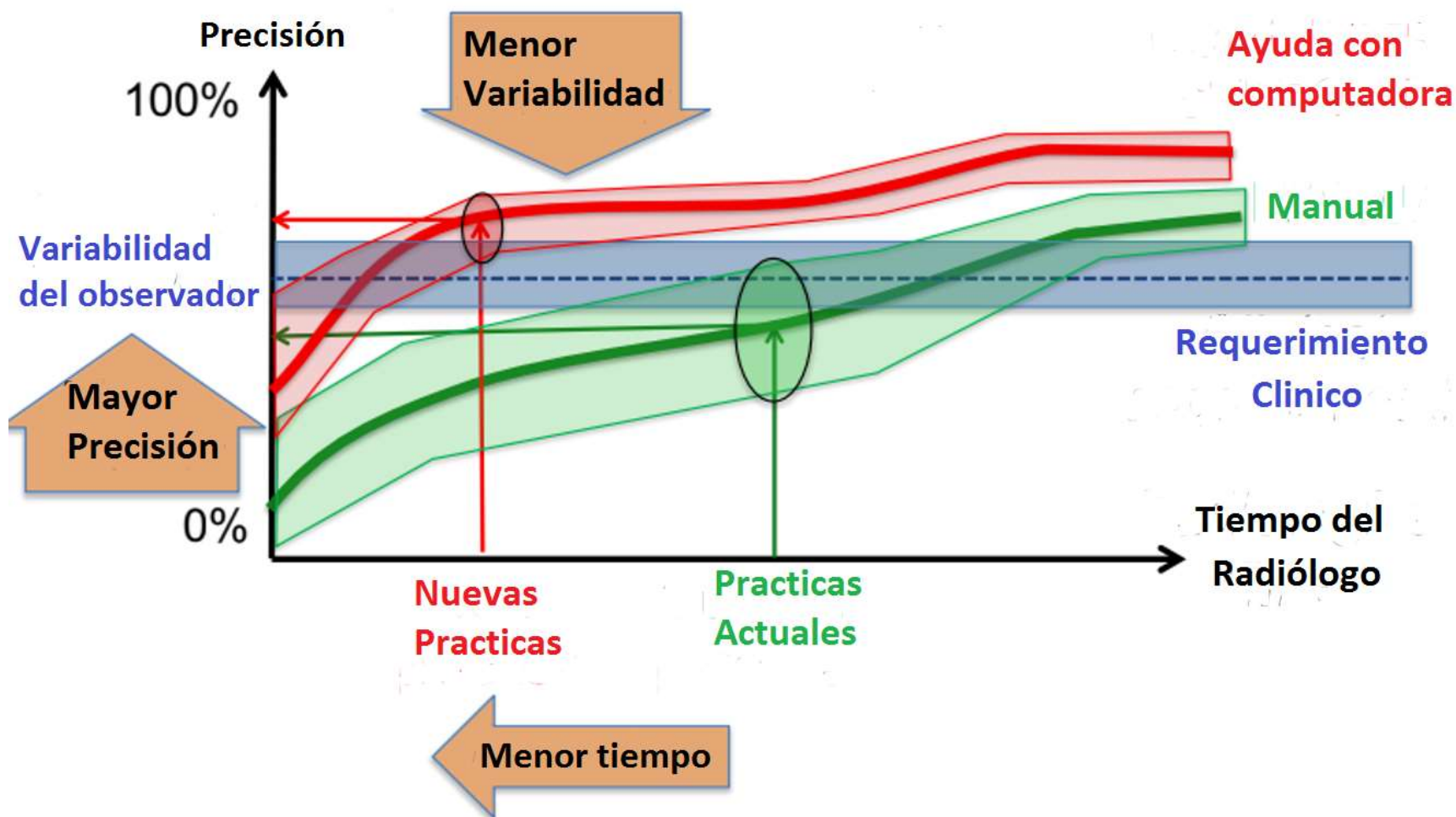


Fig. 4 Precisión y variabilidad del diagnóstico. El uso de herramientas computacionales en el apoyo al diagnóstico médico permite disminuir la variabilidad y el tiempo de este, así como alcanzar una buena precisión.



## Objetivos generales

- Desarrollar herramientas computacionales para asistir en el diagnóstico clínico en el área de imagenología médica.

## Objetivos específicos

- Optimizar el flujo de trabajo en imagenología médica mediante la automatización de tareas como la segmentación de imágenes.
- Implementar algoritmos de Redes Neuronales Convolucionales que permitan segmentar estructuras cardiacas, a partir de imágenes adquiridas por Resonancia Magnética.
- Estudiar e implementar modificaciones en estos algoritmos para mejorar la eficiencia de la segmentación.



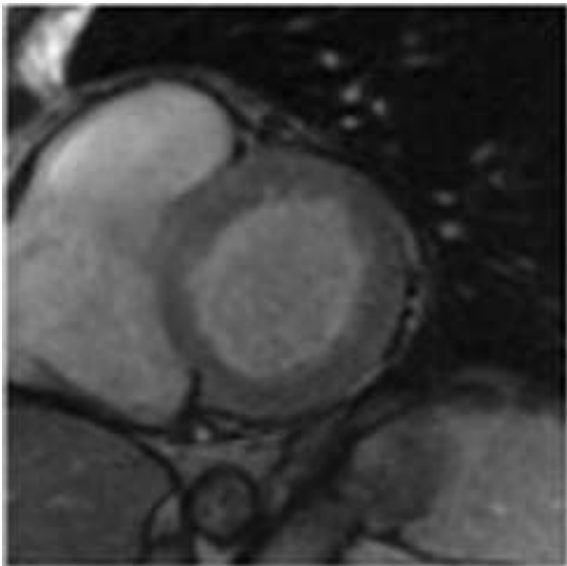
# INTRODUCCIÓN





- Distintas técnicas de segmentación han sido propuestas durante las últimas décadas, siendo estas cada vez más sofisticadas.

Métodos de umbralización. N. Otsu et. al. 1975



## Modelos activos de apariencia

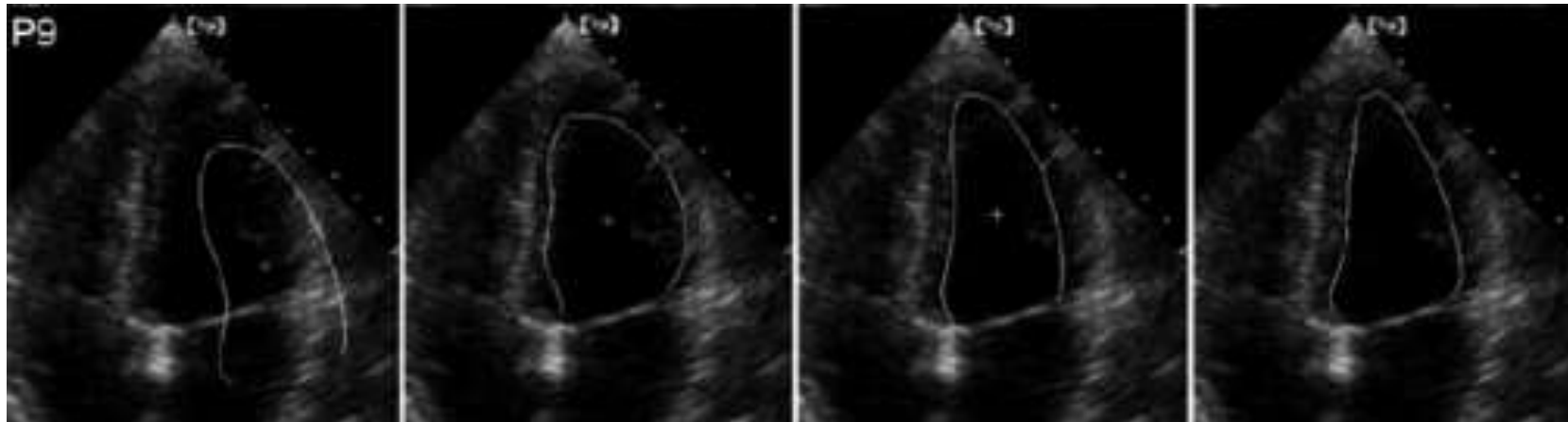


Fig. 5 Método de segmentación implementado por J. Bosch et al 2002

## Contornos activos

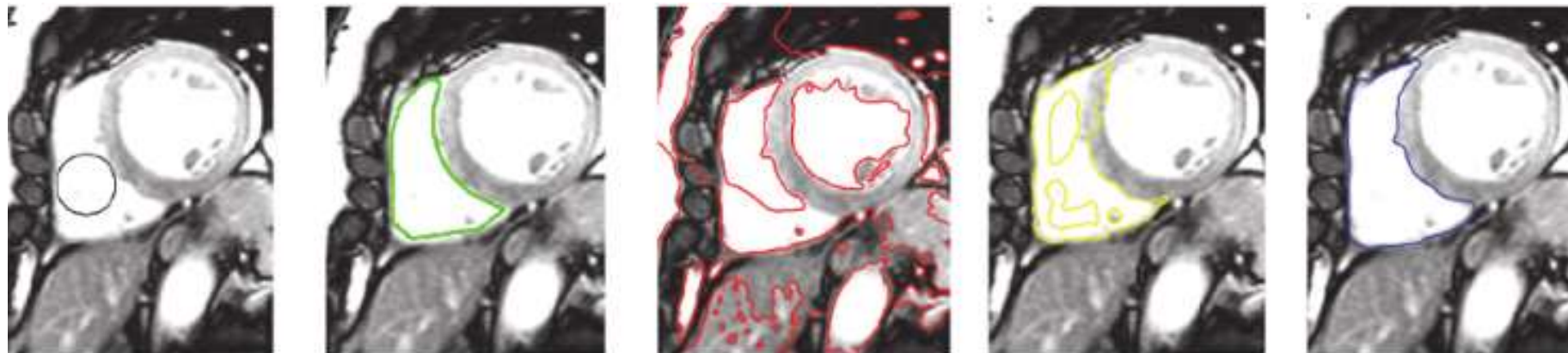


Fig. 6 Método de segmentación implementado por Shafiullah *et.al* 2017.

# Inteligencia artificial

- La Inteligencia artificial es una rama de las ciencias de la computación que busca desarrollar modelos que permiten llevar a cabo tareas similares a las del ser humano.
- Aprendizaje de máquina, consiste en el desarrollo de modelos estadísticos basados en el reconocimiento automático de patrones.
- El aprendizaje profundo se basa en la formación de representaciones complejas de los datos a partir de representaciones más sencillas.

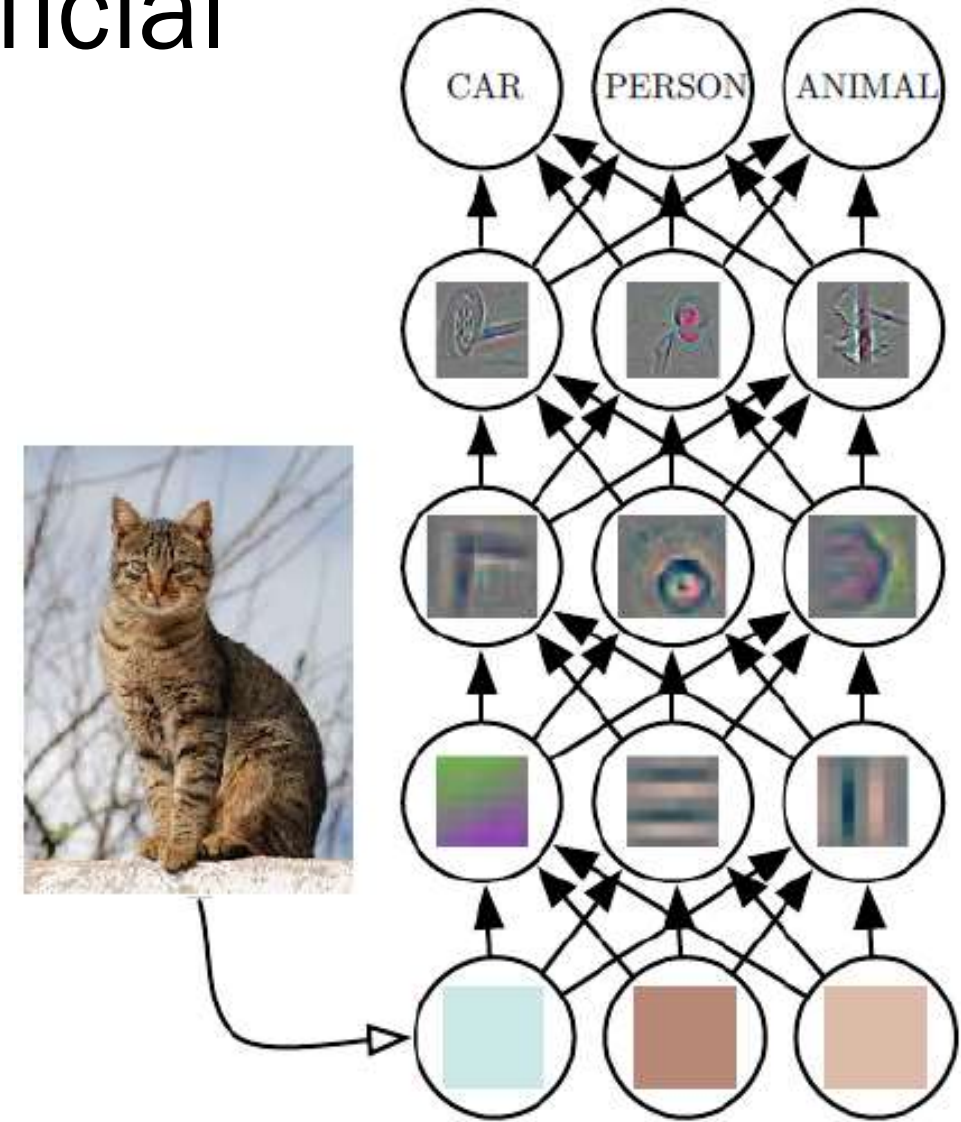


Fig. 6 Aprendizaje profundo en el reconocimiento de imágenes. I. Goodfellow et. al . (2015)

- Fei-Fei Li formó el proyecto **ImageNet**, una gran base de datos que actualmente cuenta con más de 13M de imágenes. De 2010 a 2017 llevó a cabo una competencia anual de clasificación de estas imágenes *ImageNet Large Scale Visual Recognition Challenge (ILSVRC)*

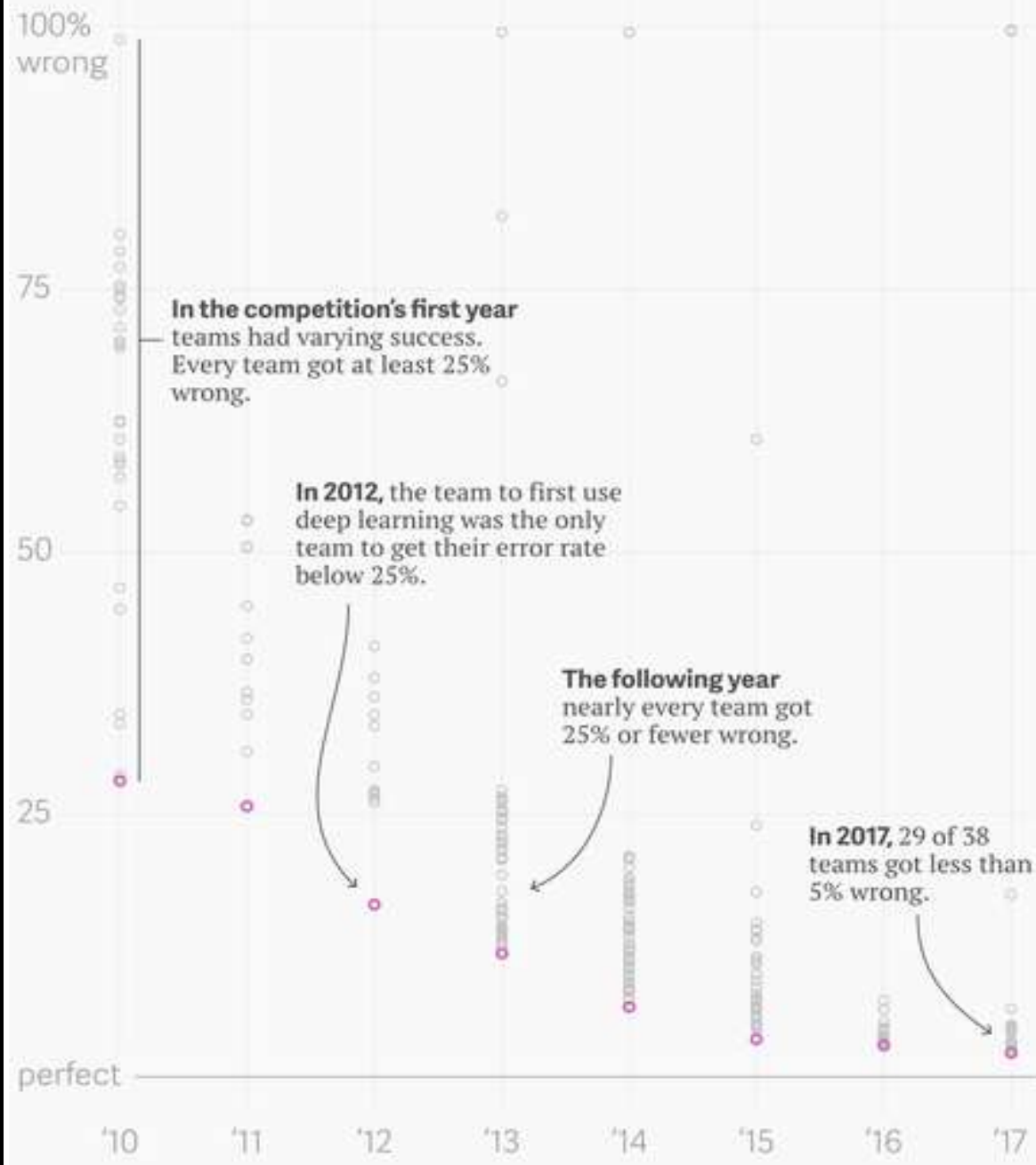


Fig .8 Resultados de la competencia ILSVRC 2010-2017

# Segmentación de imágenes

Trabajo llevado a cabo por Ronneberger Olaf et al. en 2015, en el que se propone el uso de una arquitectura tipo autocodificador, con la finalidad de llevar a cabo la segmentación en imágenes de microscopia por diferencia de fase.

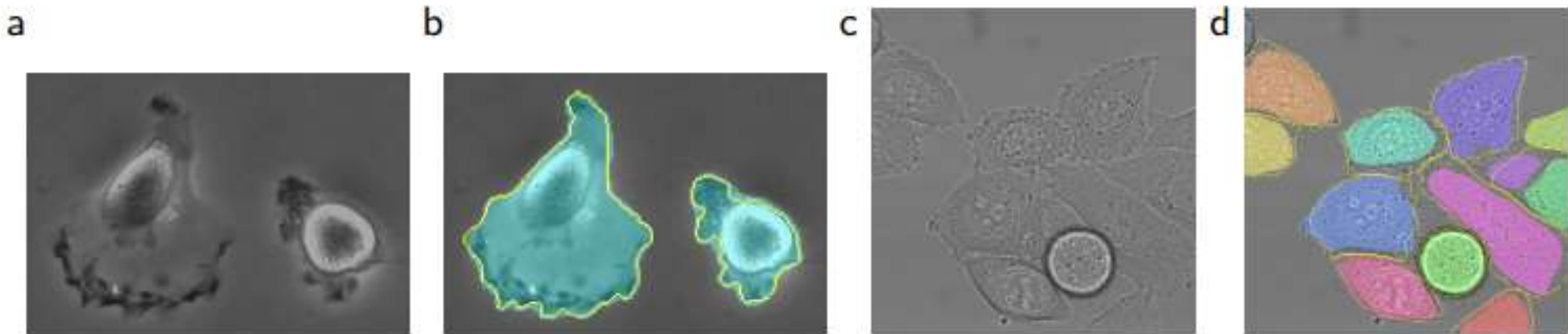


Fig .9 Segmentación en imágenes de microscopia por diferencia de fase.



## Segmentación en ecocardiografía 2D

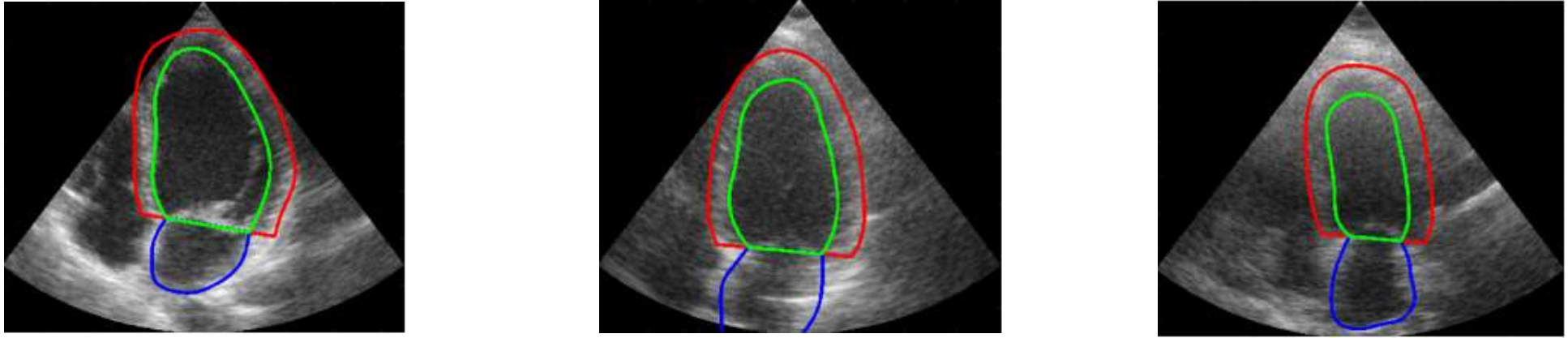


Fig .10 Trabajo llevado a cabo por Sarah Leclerc et. al. en 2019.

Clasificación de imágenes de fondo de ojo para determinar la progresión de la enfermedad de retinopatía diabética.

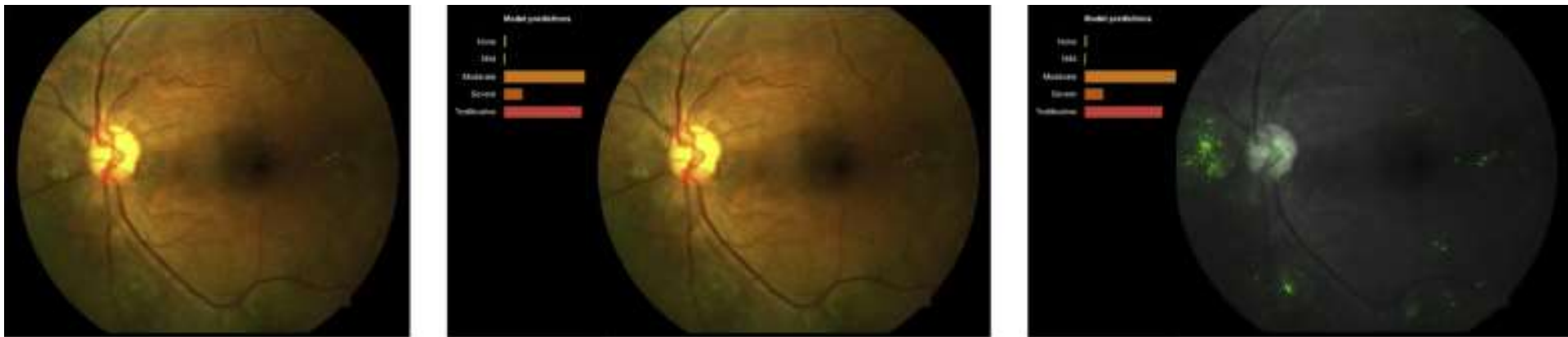
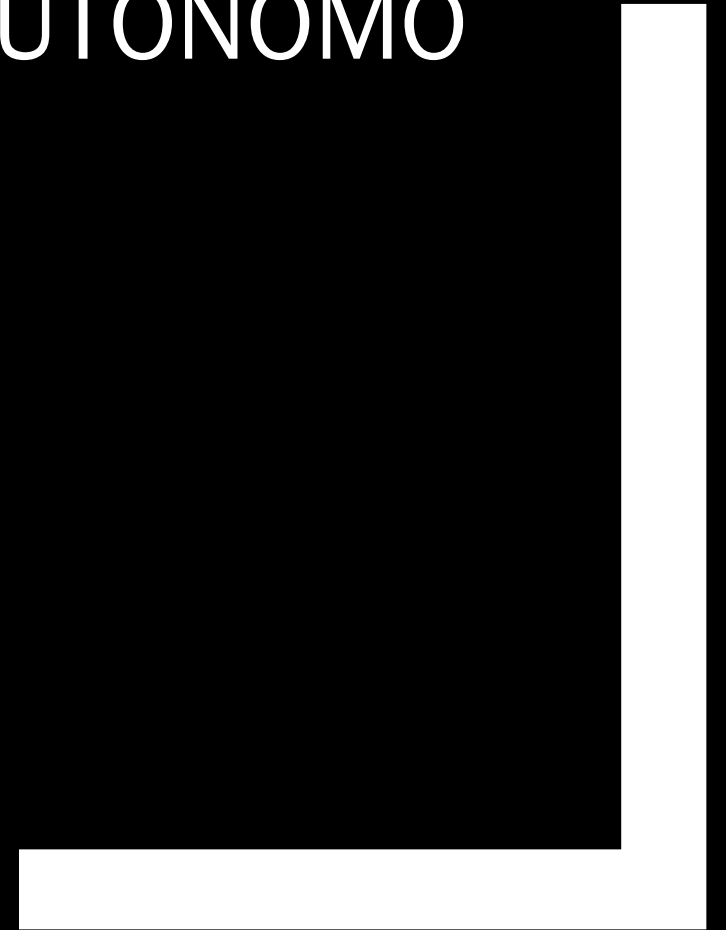


Fig .11 . Desarrollado por un equipo de Google y presentado en abril de 2019 en la revista científica *Ophthalmology*.

# ALGORITMO DE APRENDIZAJE AUTÓNOMO

## LOS DATOS





Las imágenes junto con las segmentaciones fueron tomadas de un conjunto de datos de uso libre, el cual corresponde a volúmenes de Resonancia Magnética Cardiovascular adquiridas durante una práctica clínica en el Hospital de Niños de Boston, MA,USA.

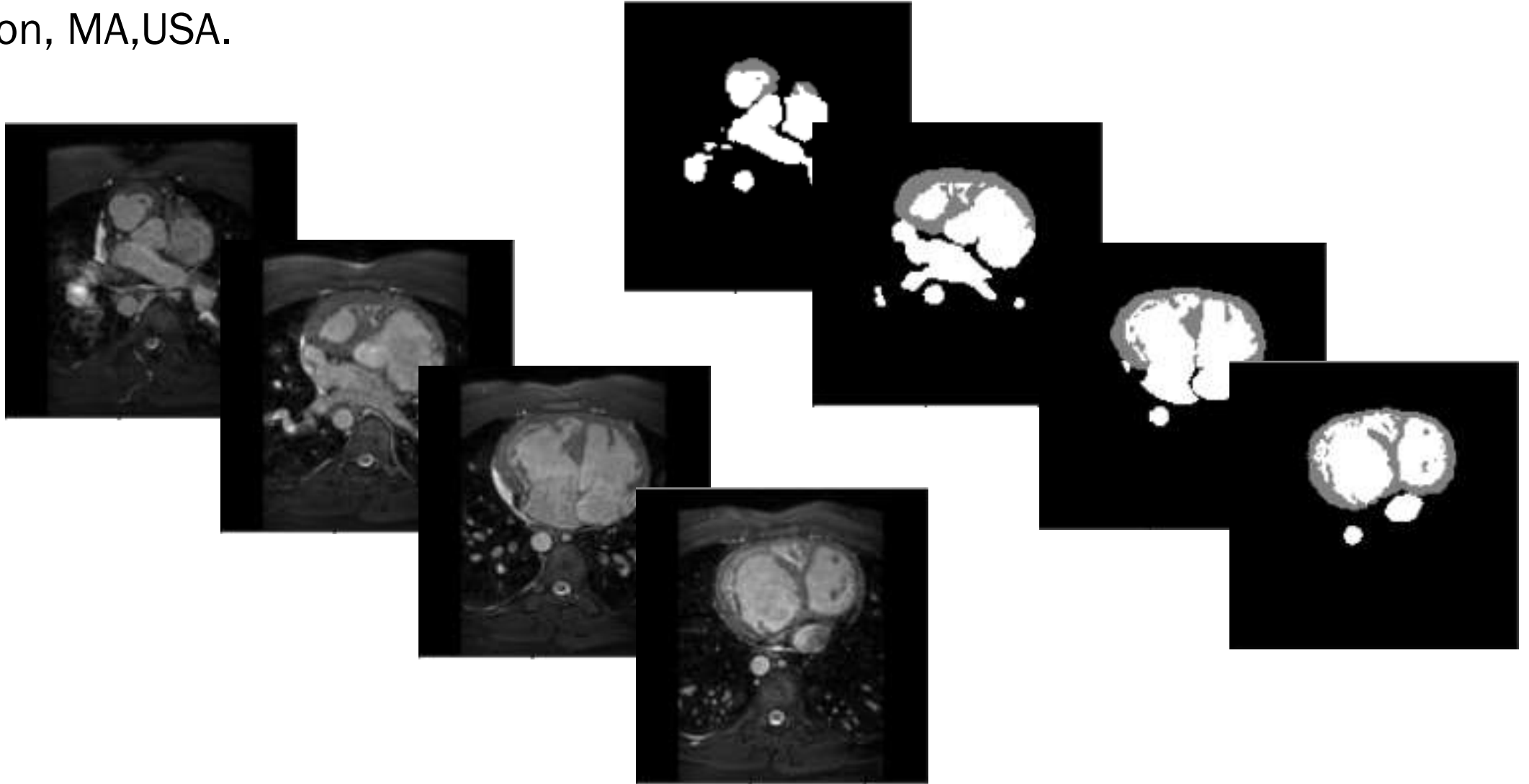
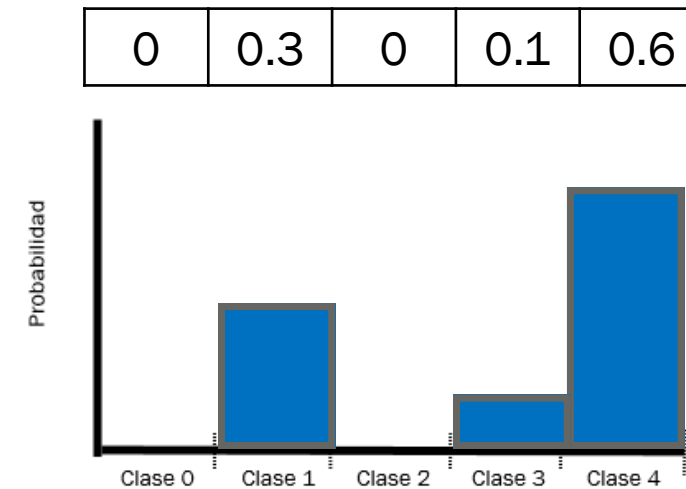
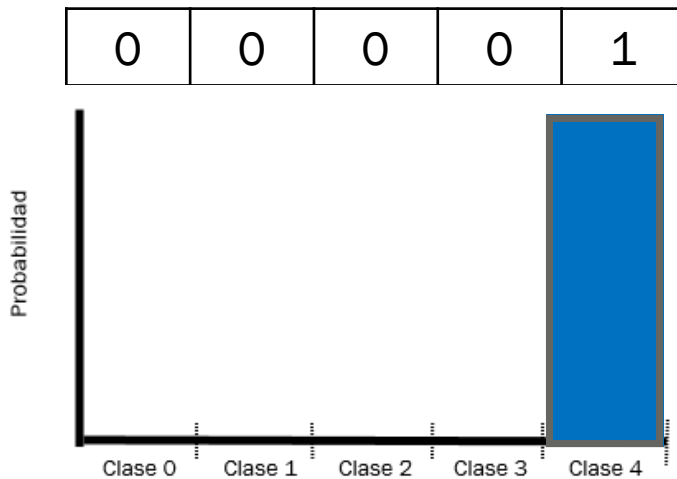
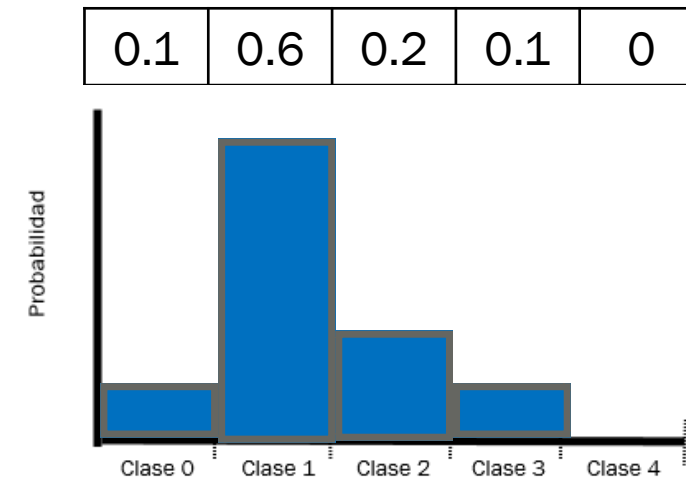
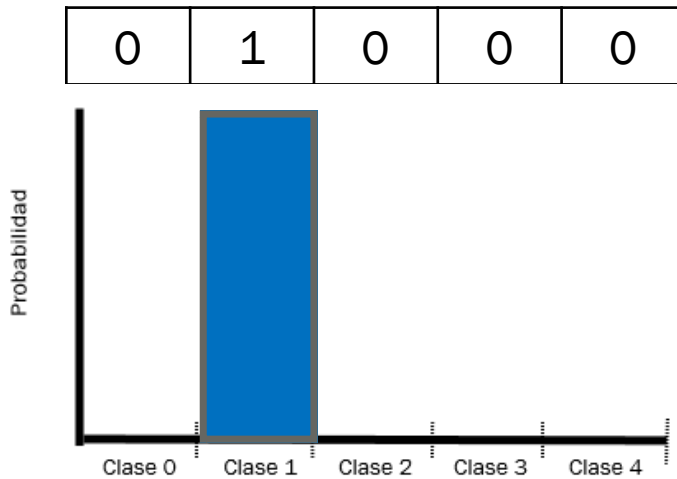


Fig. 12 Imágenes en corte transversal obtenidas de los volúmenes de la base de datos

# Distribución categórica



Distribución empírica

Distribución del modelo

Fig. 13 Imágenes en corte transversal obtenidas de los volúmenes de la base de datos

# Mapas de Segmentación

En este caso se tienen tres clases y para generar esta distribución de probabilidad los mapas de segmentación se separan en tres canales binarios donde el valor de uno o cero asigna la probabilidad de pertenecer a esa clase en particular.

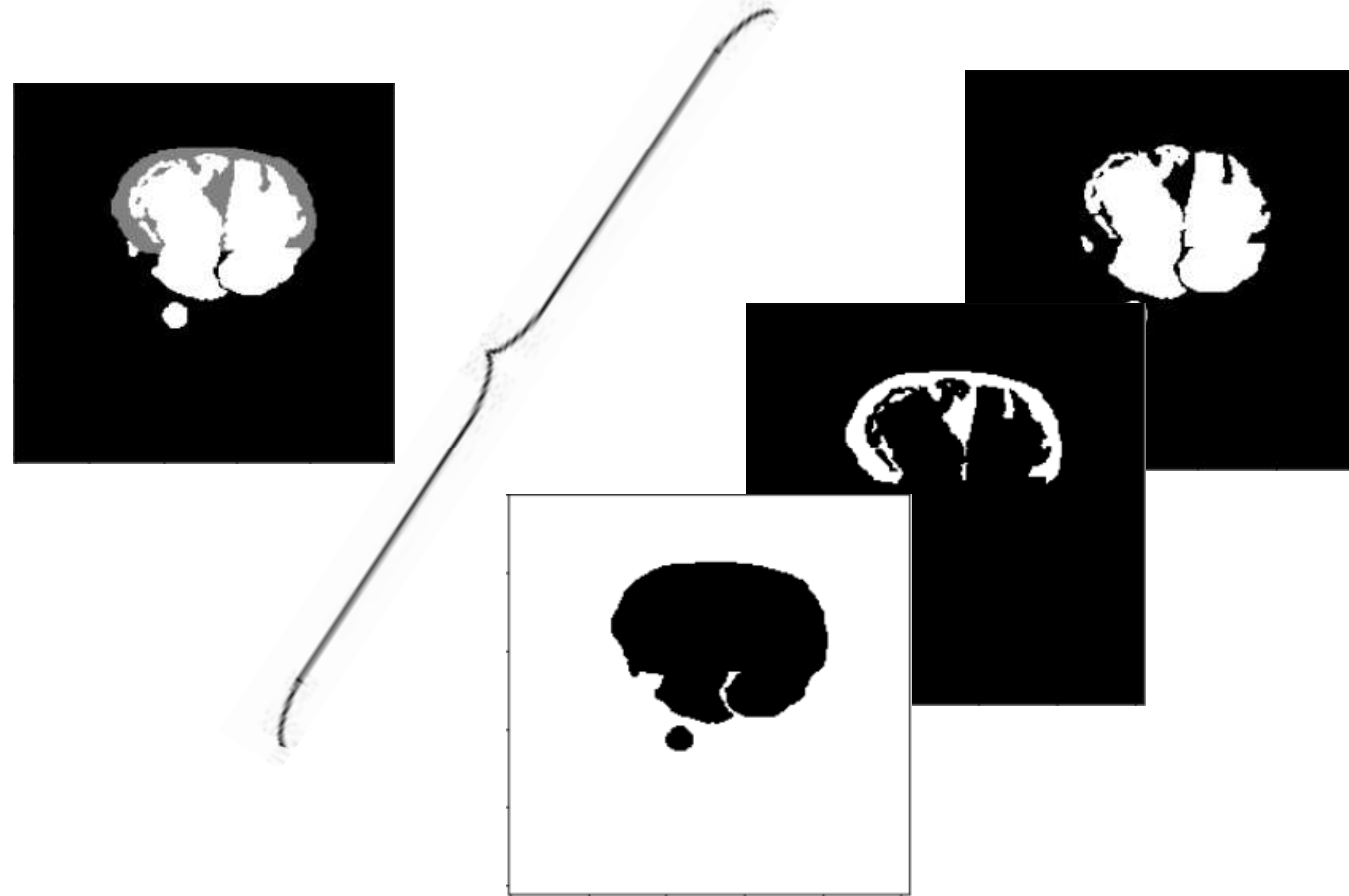


Fig. 14 Separación del mapa de segmentación en canales.

# Partición de los datos

Se llevó a cabo una partición de los datos para el entrenamiento y posterior evaluación.

- Numero de imágenes de entrenamiento: 1535
- Numero de imágenes de validación: 389
- Numero de imágenes de evaluación: 212

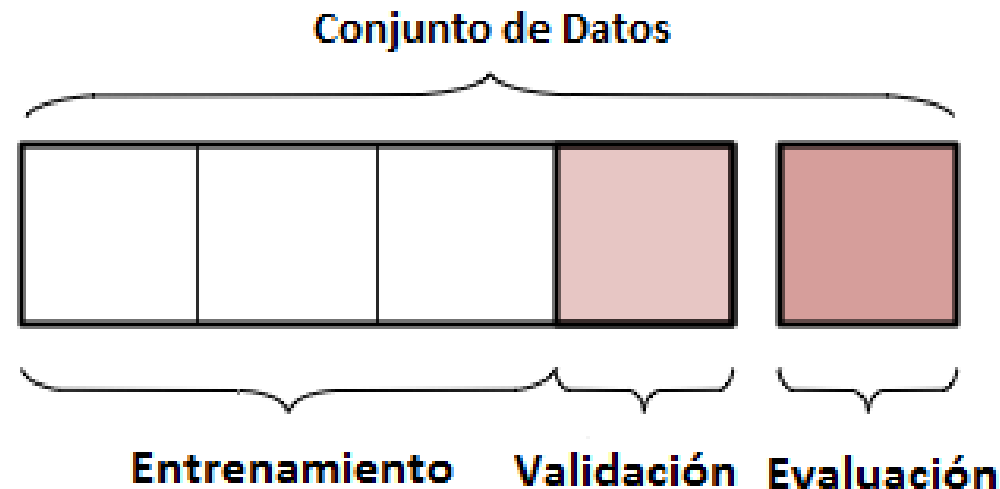
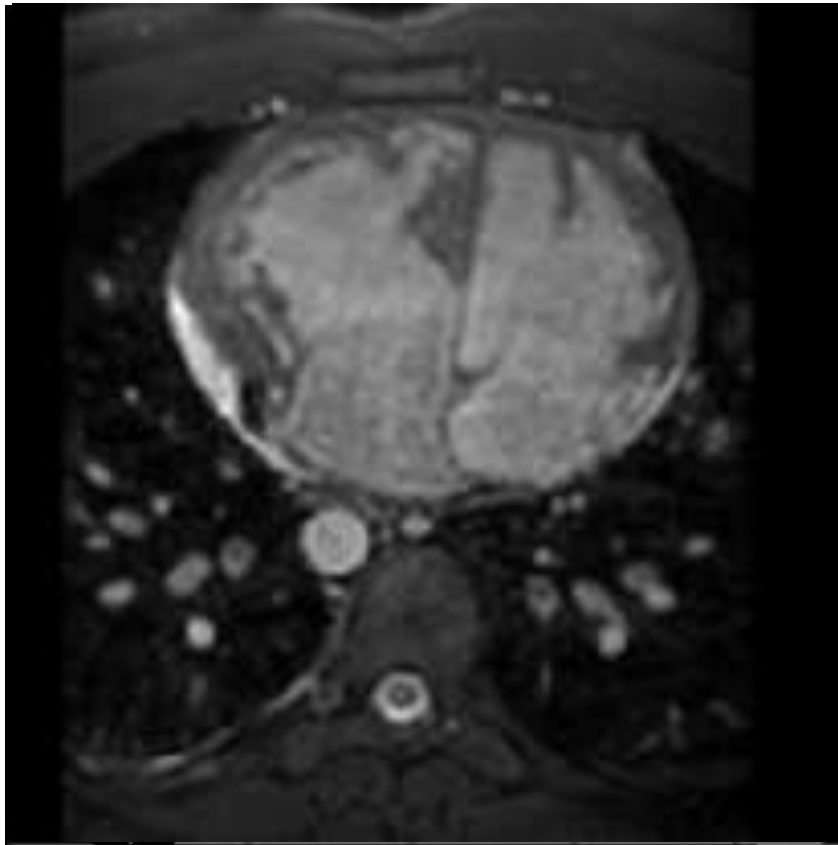


Fig. 15 Representación partición de los datos.

# Aumento de Datos

El aumento de datos se realiza para que el modelo sea más robusto a variaciones que puedan presentarse en los datos de entrada y mejorar la capacidad de generalización de este.



Rango de Rotación =  $10^{\circ}$

Rango de desplazamiento = 0.05

Rango de Re-escalamiento = 0.15

Fig. 15 Representación aumento de los datos.

ALGORITMO DE APRENDIZAJE AUTÓNOMO

ARQUITECTURA



# U-NET

Arquitectura tipo autocodificador, propuesta en 2015 para la segmentación de imágenes biomédicas.

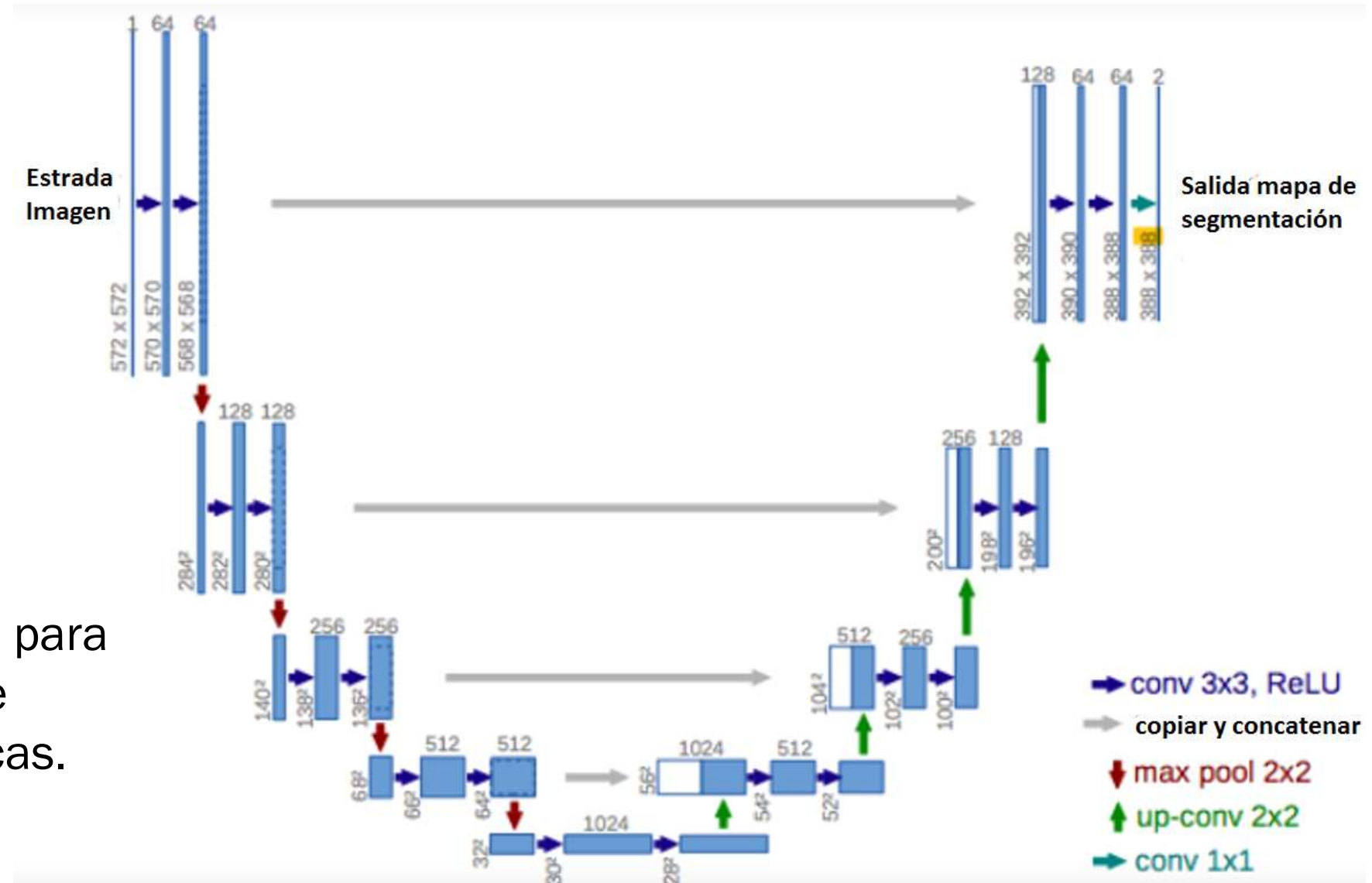
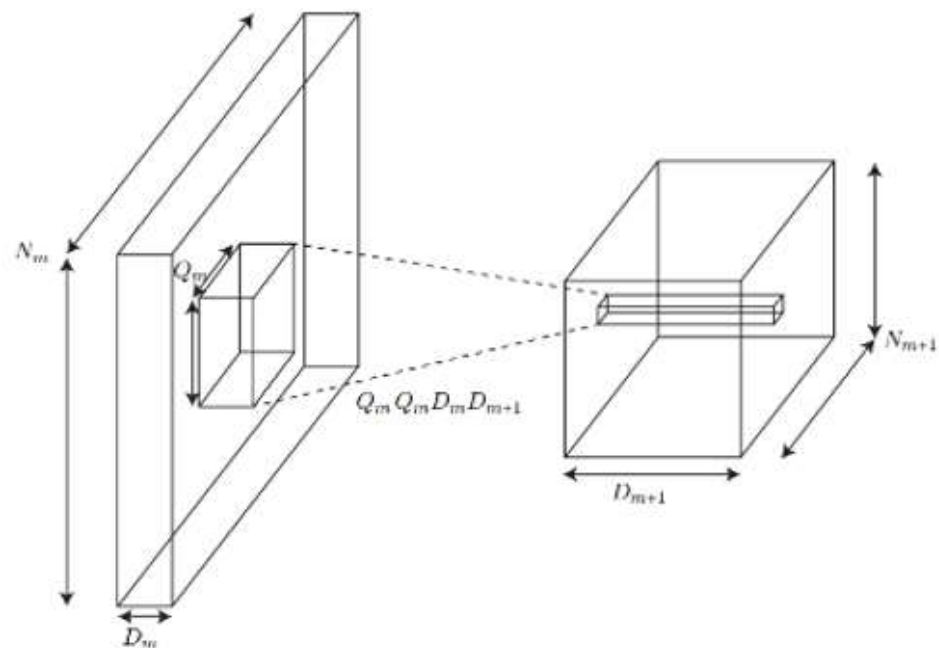


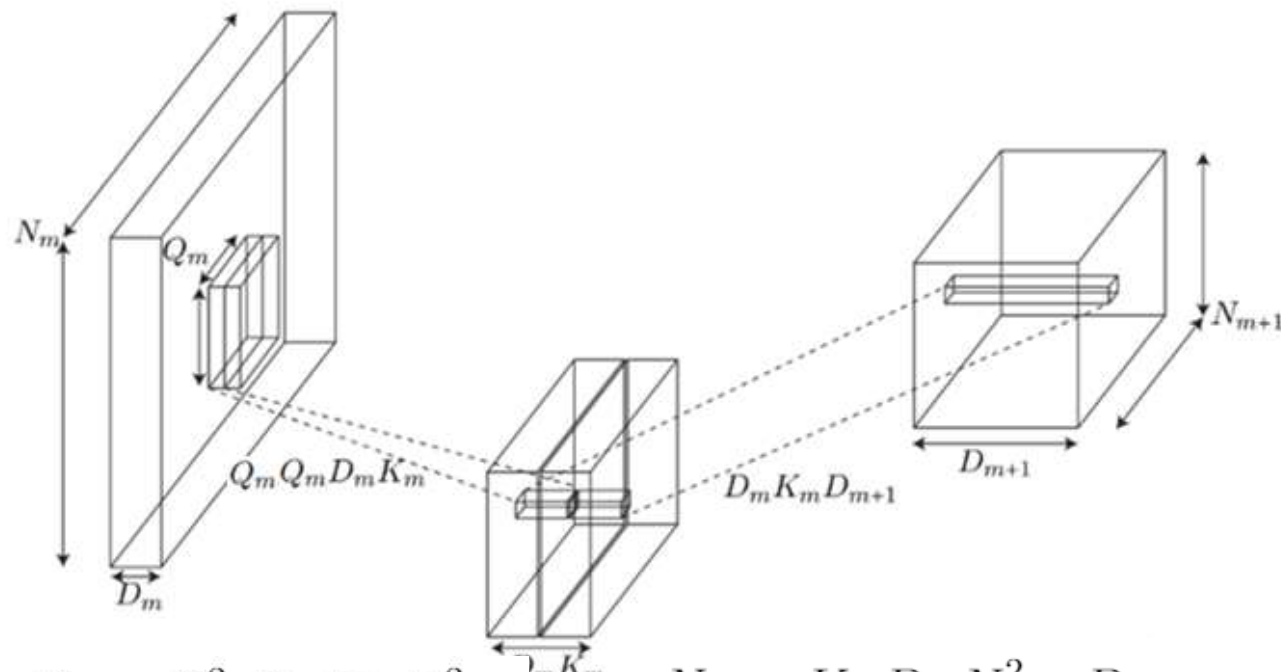
Fig. 16 Arquitectura del trabajo realizado por Ronneberger Olaf et al. en 2015.



# Comparación entre operaciones



$$N = Q_m^2 D_m D_{m+1} N_{m+1}^2$$



$$N_D = Q_m^2 D_m K_m N_{m+1}^2 \quad N_P = K_m D_m N_{m+1}^2 D_{m+1}$$

$$N_S = N_D + N_P = (Q_m^2 + D_{m+1}) D_m N_{m+1}^2 K_m$$

$$\frac{N_{mul.conv.separable}}{N_{mul.conv.convencional}} = \frac{(Q_m^2 + D_{m+1}) D_m N_{m+1}^2}{Q_m^2 D_m D_{m+1} N_{m+1}^2} = K_m \left( \frac{1}{Q_m^2} + \frac{1}{D_{m+1}} \right)$$

Fig. 17 Comparación entre la operación de convolución convencional y la convolución separable

1. Laurent Sifre. Rigid-Motion Scattering For Image Classification. PhD thesis, Ecole Polytechnique, CMAP, 2014.
2. François Chollet. Xception: Deep Learning with Depthwise Separable Convolutions 2016

# Modelos

Fue posible disminuir el número de parámetros empleados en una razón de  $\sim 5$  al emplear la operación de convolución separable.

Modelo	Parámetros	Imagen de entrada	Imagen de Salida
U-net 1	31,125,987	(256,256,1)	(256,256,3)
U-net 2	6,030,156	(256,256,1)	(256,256,3)

Tabla. 1 Modelos implementados. U-net 1) Uso de la operación de convolución convencional.  
U-net 2) Uso de la operación de convolución separable.

# Softmax

Como se desea trabajar con probabilidades, es necesario que los valores de salida de la red cumplan propiedades particulares

$$p(y_i | x) \leq 0$$

$$\sum_i^N p(y_i | x) = 1$$

Esto se puede lograr al aplicar la función  $\text{softmax}(z_i)$  a los valores de salida del modelo. Donde  $z_i$  representa el valor de la probabilidad (sin normalizar) de pertenecer a la categoría  $i$ -ésima .

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_i e^{z_i}}$$

# ALGORITMO DE APRENDIZAJE AUTÓNOMO

## Entrenamiento



# Aprendizaje supervisado

- La **Estimación por Máxima Verosimilitud** (EMV) es un método desarrollado en el área de estadística, que permite estimar los parámetros de una distribución probabilística, lo cual puede ser llevado a cabo mediante la maximización de una **función de verosimilitud**.
- En este tipo de aprendizaje se cuenta con un conjunto de datos u observaciones  $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$  y con un conjunto de etiquetas  $\mathbf{Y} = \{y_1, y_2, \dots, y_n\}$ .
- $\hat{p}_{data}(y|x)$ : Modelo empírico el cual es el resultado del conjunto de observaciones con el que se cuenta.
- $p_{model}(y|x; \theta)$ : Modelo predefinido y que esta parametrizado por  $\theta$ .

# Función de costo

El principio de máxima verosimilitud proporciona una forma de cuantificar la diferencia entre la distribución empírica generada por los datos observados y la distribución aprendida por nuestro modelo

$$CCE = - \sum_c \sum_i \hat{p}_{data}(y_i|x_i) [\log p_{model}(y_i|x_i)]$$

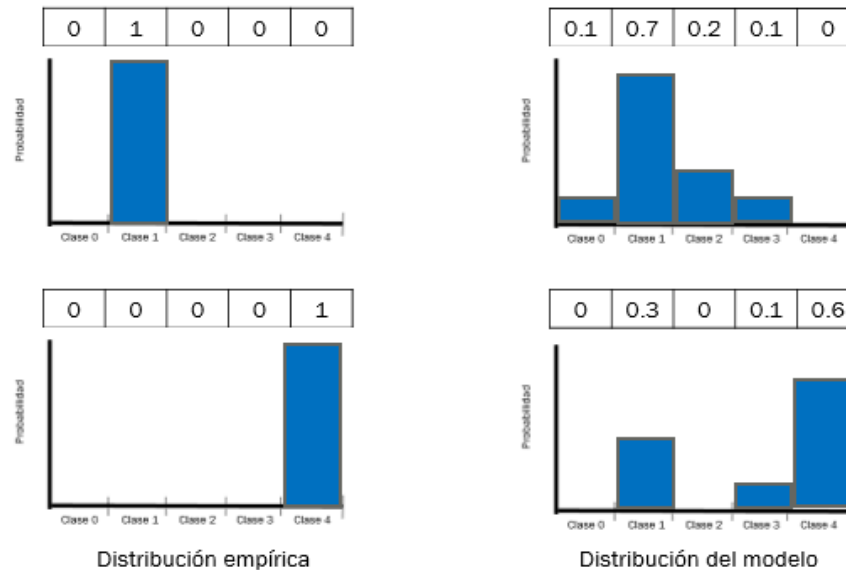


Fig. 13 Imágenes en corte transversal obtenidas de los volúmenes de la base de datos

# Coeficiente de Dice

Es una medida de que tanto se sobreponen dos estructuras, lo cual permite medir la similitud entre el mapa de segmentación generado por el modelo y aquel generado por un médico experto.

$$DC = \frac{2 \sum_N^i p_{model}(y_i|x_i) \hat{p}_{data}(y_i|x_i)}{\sum_N^i p_{model}(y_i|x_i)^2 + \sum_N^i \hat{p}_{data}(y_i|x_i)^2}$$

$$DL = 1 - DC$$

$$LOSS = CCE + DL$$

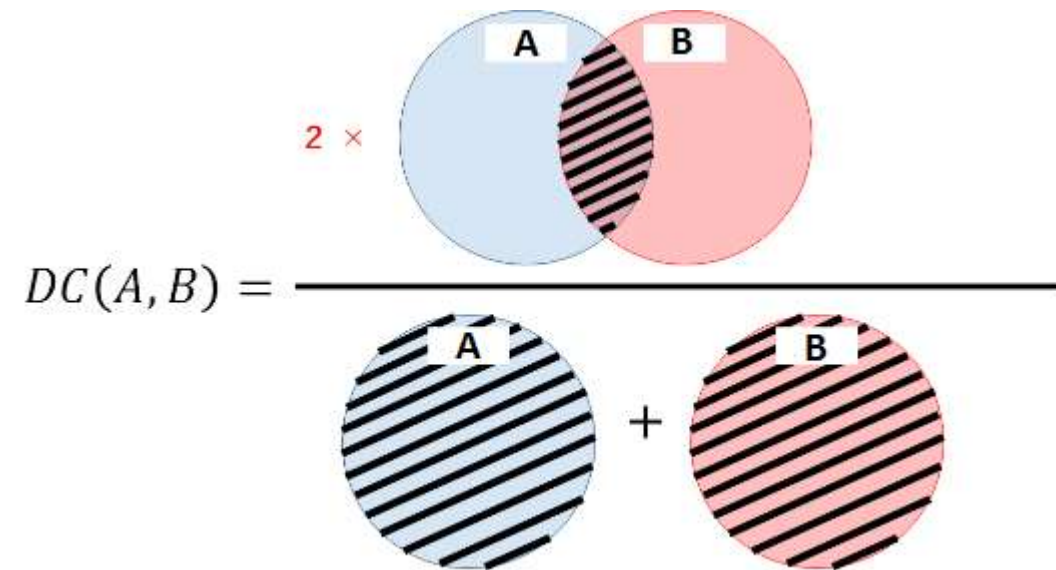


Fig. 20 Representación del calculo del coeficiente de Dice.



# Método de Optimización

En el caso de que la ecuación anterior describa una función convexa, y que esta sea diferenciable en  $\theta$ , encontrar un punto critico se reduce al criterio:

$$\frac{\partial L(X; \theta)}{\partial \theta_1} = \frac{\partial L(X; \theta)}{\partial \theta_2} = \dots = \frac{\partial L(X; \theta)}{\partial \theta_n} = 0$$

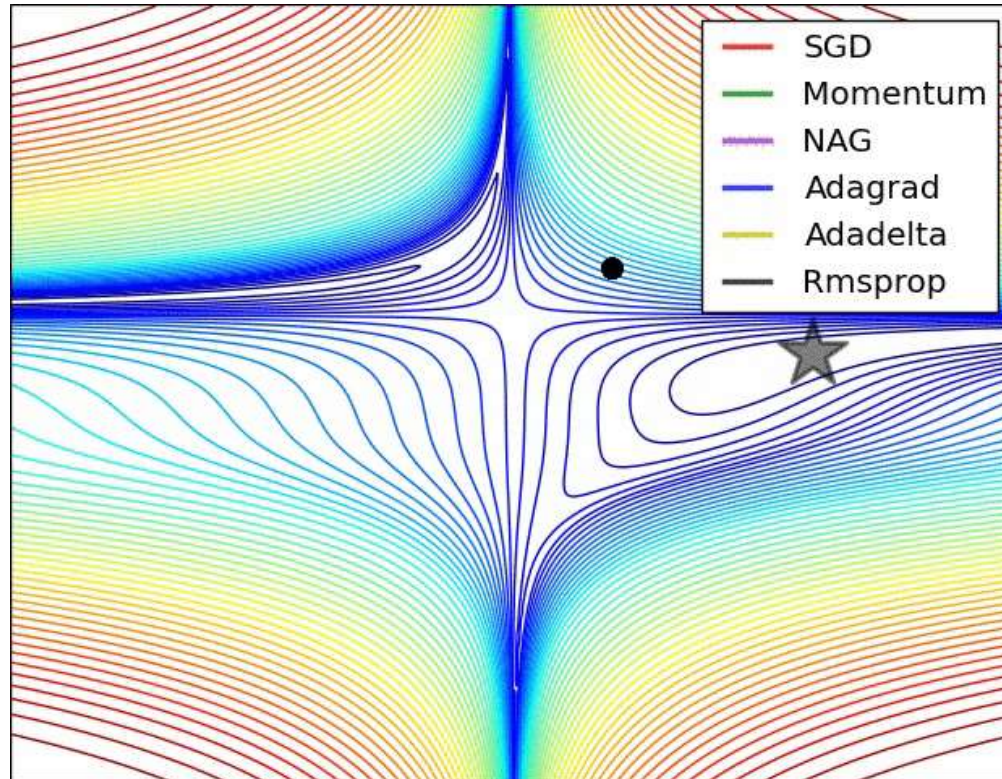
Para modelos con una gran cantidad de parámetros se recurre a métodos iterativos como **gradiente descendente**

$$\theta_{i+1} = \theta_i - \varepsilon \nabla L(X; \theta)$$

# Estrategia de Optimización

La estrategia de optimización empleada fue Adam<sup>1</sup> ('Adaptative moments'), una combinación de 'Momentum' y 'RMSProp'.

$$\mathbf{g} \leftarrow \frac{1}{m} \nabla_{\boldsymbol{\theta}} \sum_i L(p_{\text{model}}(y_i|x_i; \boldsymbol{\theta}), \hat{p}_{\text{data}}(y_i|x_i))$$



$$\mathbf{s} \leftarrow \rho_1 \mathbf{s} + (1 - \rho_1) \mathbf{g}$$

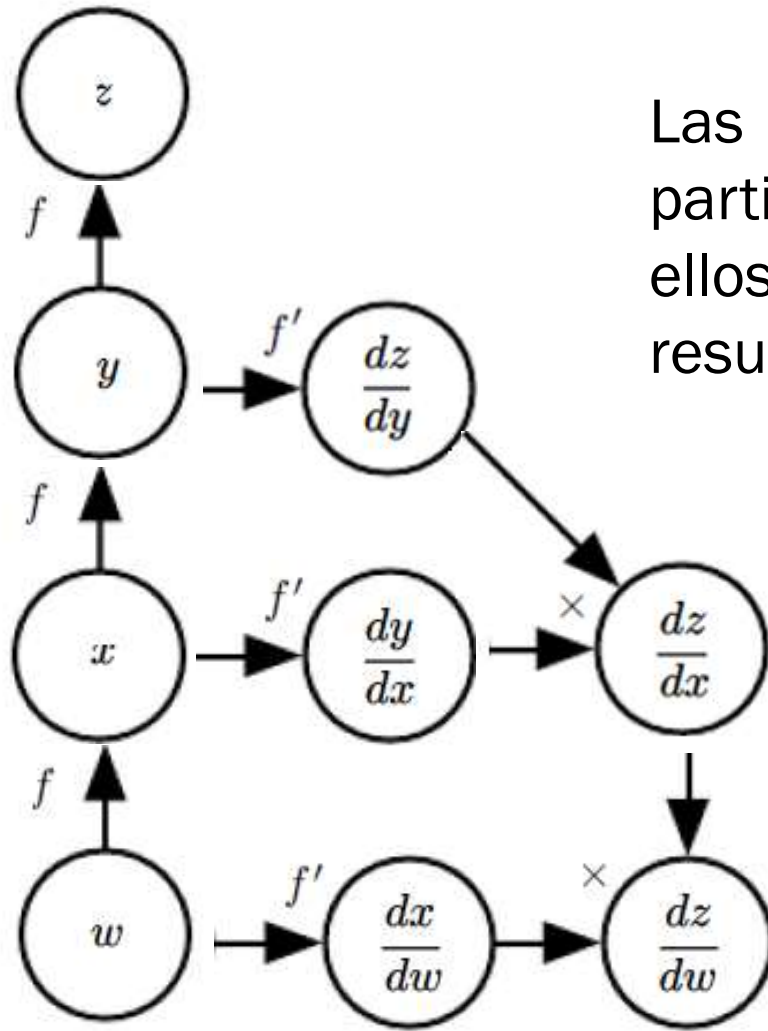
$$\mathbf{r} \leftarrow \rho_2 \mathbf{r} + (1 - \rho_2) \mathbf{g} \odot \mathbf{g}$$

$$\hat{\mathbf{s}} \leftarrow \frac{\mathbf{s}}{1 - \rho_1^t} \quad \hat{\mathbf{r}} \leftarrow \frac{\mathbf{r}}{1 - \rho_2^t}$$

$$\Delta \boldsymbol{\theta} \leftarrow -\varepsilon \frac{\hat{\mathbf{s}}}{\sqrt{\hat{\mathbf{r}} - \delta}}$$

$$\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} + \Delta \boldsymbol{\theta}$$

# Algoritmo de retro-propagación



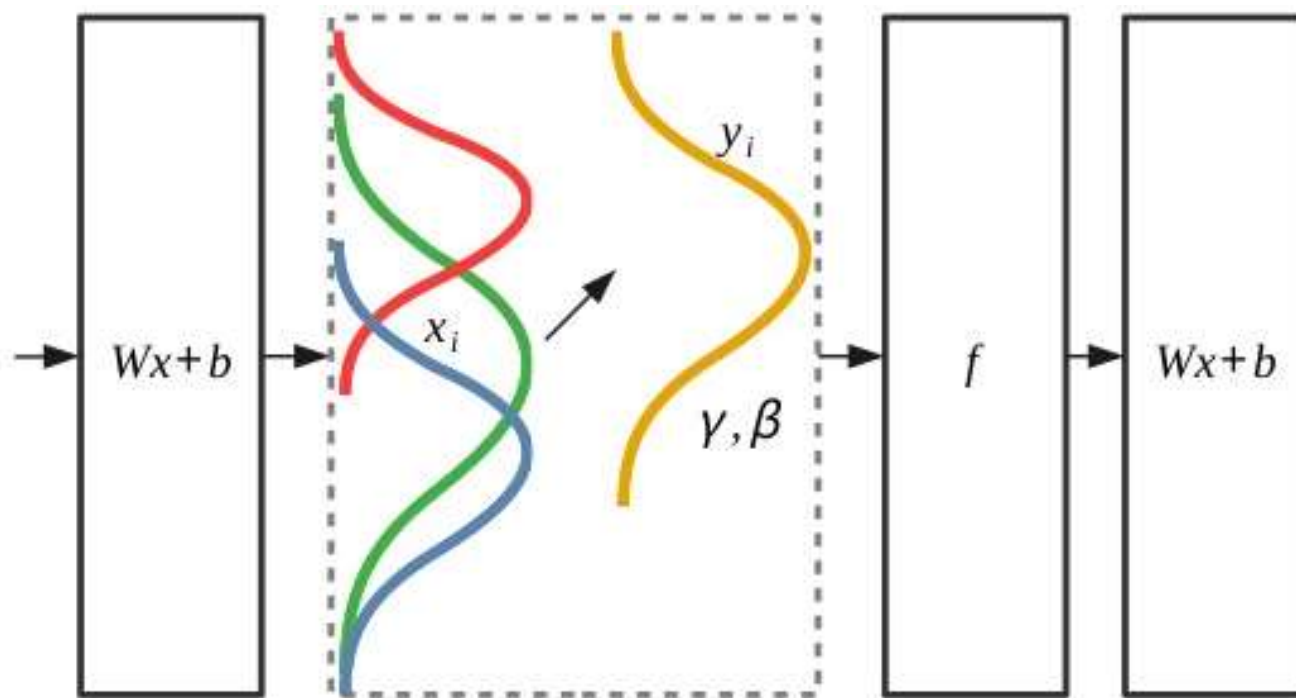
Las unidades operacionales de una arquitectura en particular pueden visualizarse como nodos, cada uno de ellos con parámetros modificables que influyen en el resultado final.

$$\frac{\partial u^{(n)}}{\partial u^{(j)}} = \sum_{i: j \in Pa(u^{(i)})} \frac{\partial u^{(n)}}{\partial u^{(i)}} \frac{\partial u^{(i)}}{\partial u^{(j)}}$$

Fig. 21 Esquema del flujo de datos en el modelo.

# Normalización por lote

La modificación de los parámetros en las primeras capas puede tener un gran efecto en la modificación de la distribución de los valores de entrada que tendrán las capa siguientes y en particular, las más profundas (**internal covariate shift**<sup>1</sup>).



$$\hat{x}_i \leftarrow \frac{x_i - \mu}{\sqrt{\sigma^2 + \epsilon}}$$

$$y_i \leftarrow \mathbf{BN}_{\gamma, \beta}(\hat{x}_i) = \gamma \hat{x}_i + \beta$$

Fig. 19 Representación del efecto de desplazamiento interno de la covariable.

1. Sergey Ioffe y Christian Szegedy, 2015, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift

# Regularización

Un modelo con una gran cantidad de parámetros tiene la **capacidad** de aprender distintas distribuciones, sin embargo, muchas veces esto puede resultar en un sobreajuste del modelo a los datos de entrenamiento<sup>1</sup>.

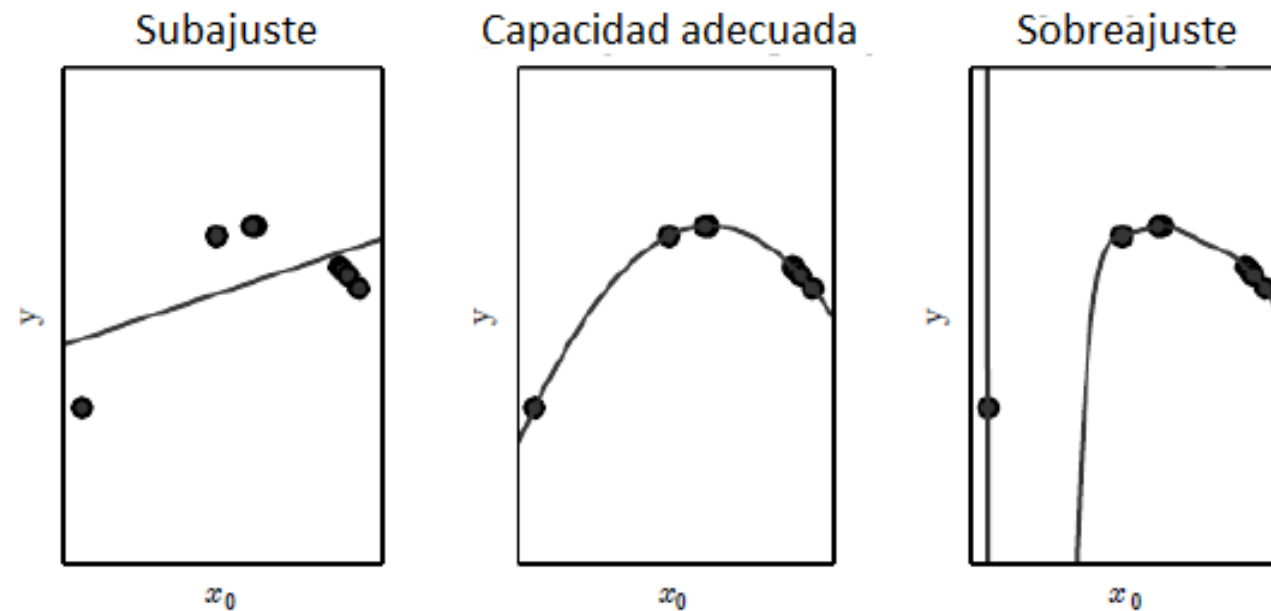


Fig. 18 Representación de la influencia de la capacidad del modelo en el ajuste de los datos.

Una forma de corregir esto es usando métodos de regularización como **2D Spatial Dropout**

1. J. Tompson et.al 2014 Efficient Object Localization Using Convolutional Networks.



# RESULTADOS



# Desempeño

Se obtuvieron resultados cuantitativos al estudiar el desempeño del modelo sobre las 212 imágenes de conjunto evaluación.

Modelo	Métrica	Fondo	Miocardio	Volumen Sanguíneo	Promedio
U-net 1	Dice	0.990	0.748	0.747	0.829
	CCE	0.018	0.049	0.070	0.042
U-net 2	Dice	0.992	0.793	0.751	0.845
	CCE	0.007	0.035	0.056	0.035

Tabla 2. Resultados cuantitativos del desempeño del modelo.



# Comparación cualitativa

**Columna izquierda)** Mascara de segmentación original.

**Columna Central)** Mascara de segmentación U-net 1.

**Columna Derecha)** Mascara de segmentación U-net 2.

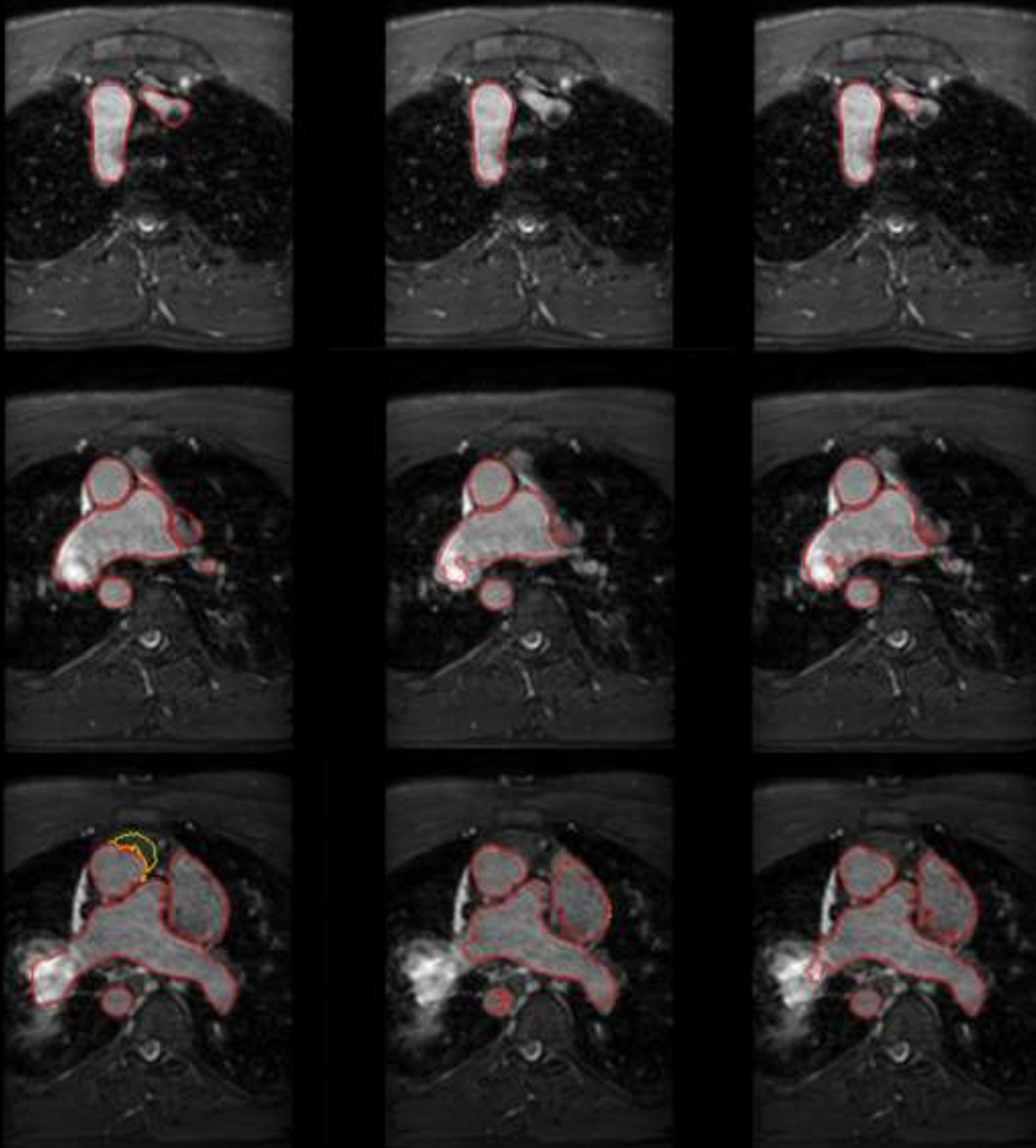


Fig 22.1 Comparaciones cualitativas del desempeño del modelo.

# Comparación cualitativa

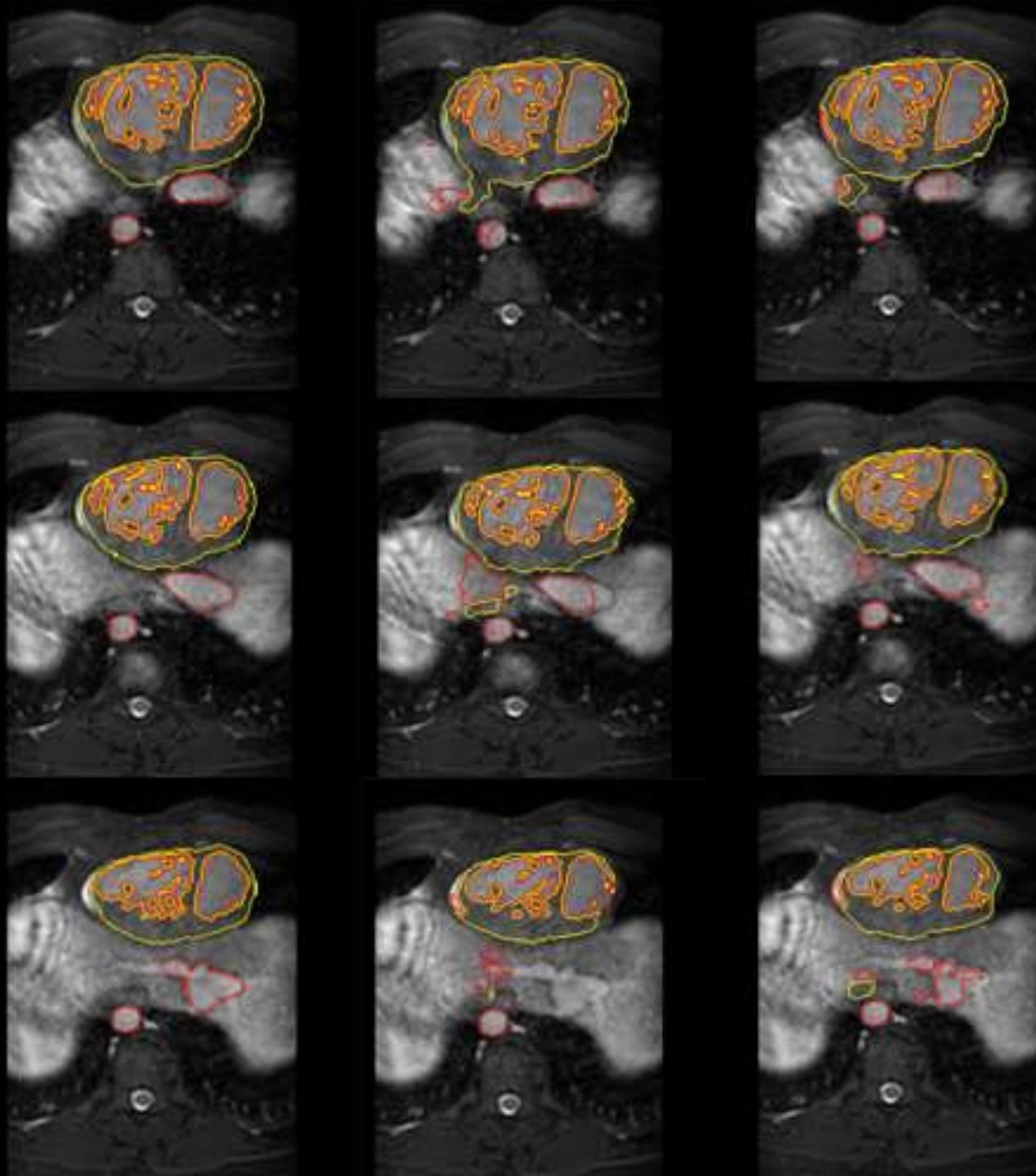
**Columna izquierda)** Mascara de segmentación original.

**Columna Central)** Mascara de segmentación U-net 1.

**Columna Derecha)** Mascara de segmentación U-net 2.

Fig 22.2 Comparaciones cualitativas del desempeño del modelo.

# Comparación cualitativa



**Columna izquierda)** Mascara de segmentación original.

**Columna Central)** Mascara de segmentación U-net 1.

**Columna Derecha)** Mascara de segmentación U-net 2.

Fig 22.2 Comparaciones cualitativas del desempeño del modelo.



# Conclusiones

- Fue posible implementar la arquitectura tipo auto-encoder propuesta en el trabajo de *Ronneberger et al.* en 2015, así como compararla con una nueva versión desarrollada a lo largo de este trabajo.
- La sustitución de la operación de convolución convencional por convolución separable permite obtener un modelo que alcanza un desempeño comparable e incluso superior en esta tarea, pero disminuyendo significativamente (en una razón de  $\sim 5$ ) la cantidad de parámetros empleados por la red, lo que puede atribuirse a un uso más óptimo estos.
- Las herramientas desarrolladas en este trabajo presentan una aplicación con un alto potencial en el área de segmentación en imágenes médicas.

# Trabajo a futuro

Algunas otras técnicas que permiten mejorar el desempeño de la red:

- La optimización de hiperparametros del modelo, como el valor de la tasa de aprendizaje inicial, el tamaño de lote de entrenamiento, el porcentaje de regularización, etc.
- Incorporar factores adicionales a la función de costo que permitan tomar en cuenta el desbalance de clases existente.

En un futuro, seria ideal el poder usar estas técnicas en combinación, con el propósito de que el desempeño sea lo suficientemente bueno como para que el algoritmo sea incorporado.

Gracias por su atención