

DESARROLLO DE UNA HERRAMIENTA PARA LA DETECCIÓN AUTOMÁTICA DEL PLANO VALVULAR MITRAL MEDIANTE ALGORITMOS DE DEEP LEARNING

M. Giménez Ucendo¹, D. Bermejo-Peláez¹, J.E Ortuño Fisac¹, E.R McVeigh², M.J Ledesma-Carbayo¹

¹ Biomedical Image Technologies, Universidad Politécnica de Madrid & CIBER Madrid, España.

² Department of Bioengineering, UC San Diego, La Jolla, CA, USA

Resumen

Las enfermedades cardiovasculares son la principal causa de muerte en España. Este hecho incentiva el desarrollo de herramientas automáticas de ayuda al diagnóstico de las mismas basadas en imágenes médicas. Por otro lado, una de las estructuras cardíacas con mayor importancia en el corazón es la válvula mitral. La localización y detección automática del plano que la contiene puede ayudar a la obtención de biomarcadores relacionados con la función del ventrículo izquierdo, así como apoyar en el diagnóstico y la planificación terapéutica de las diversas patologías en las que puede estar involucrada.

En este trabajo se propone un método basado en Redes Neuronales Convolucionales (CNN) para la detección y localización automática del plano valvular mitral de forma dinámica a lo largo de ciclo cardíaco a partir de secuencias de imágenes de CT cardíaco. El método propuesto se ha entrenado y validado con 96 frames correspondientes a 12 pacientes diferentes.

1. Motivación

Las enfermedades cardiovasculares son la principal causa de muerte de nuestro país, representando según el Instituto Nacional de Estadística un 29,2% de los fallecimientos en el 2016. Tanto los hábitos saludables como el diagnóstico precoz de dichas enfermedades son de vital importancia para disminuir el grado de afectación. Una de las estructuras cardíacas de alta importancia para gran parte de estas enfermedades es la válvula mitral [1]. La válvula mitral es la encargada de regular el flujo sanguíneo desde la aurícula izquierda al ventrículo izquierdo, y el estudio del plano que define nos proporciona gran cantidad de información.

En primer lugar, la localización del plano mitral nos permite analizar el estado de la válvula y diagnosticar un conjunto de enfermedades cardiovasculares conocido como valvulopatías asociadas a dicha válvula. Las valvulopatías son las enfermedades propias de las válvulas del corazón. En la actualidad, la prueba principal que se utiliza para diagnosticar este tipo de patologías es el ecocardiograma, la cual es la más rápida y económica. Sin embargo, para realizar un diagnóstico profundo, y, sobre todo, para realizar una correcta y minuciosa localización de la válvula mitral de cara a la planificación de una posible intervención, es necesario el uso de otras modalidades de imagen con mayor resolución espacial como la Tomografía computarizada (CT). El tratamiento de las valvulopatías no sólo se ha desarrollado de manera notable en las últimas décadas por el continuo progreso de la tecnología protésica valvular, sino que también se ha reorientado hacia el desarrollo de intervenciones quirúrgicas reparadoras y la introducción de técnicas de intervención percutánea. En definitiva, es necesario llevar a cabo un estudio previo muy detallado de la localización y estado de la válvula mitral,

que a su vez queda definido por la observación y medidas en el entorno del plano valvular mitral que determinan el posterior tratamiento.

En segundo lugar, el plano valvular mitral no solo es útil para la localización de las válvulas para su diagnóstico y planificación terapéutica, sino que la detección del plano valvular mitral sirve para llevar a cabo multitud de medidas cuantitativas de la propia válvula, como por ejemplo medidas del diámetro valvular, área valvular, así como la obtención de otros parámetros funcionales del ventrículo izquierdo como la fracción del volumen de eyección durante el ciclo cardíaco de manera más precisa. Estas medidas tienen gran valor diagnóstico y permiten definir la estrategia de tratamiento o intervención más adecuada.

Por otro lado, a lo largo del ciclo cardíaco el corazón está en continuo movimiento. Para una correcta visualización (y posterior cuantificación) de ciertas estructuras del corazón de forma dinámica a lo largo del ciclo cardíaco, se requiere de una corrección de este movimiento. Dicha corrección de movimiento se podría llevar a cabo en un volumen reformateado según el eje largo del corazón, el cual dependería de una previa localización del plano valvular mitral

Hasta ahora, la localización del plano valvular mitral se ha realizado de forma manual, conllevando de esta forma un alto coste tanto en términos económicos como de tiempo. Es por ello que el desarrollo de una herramienta capaz de detectar el plano valvular mitral de manera automática durante el ciclo cardíaco podría favorecer la realización de medidas más precisas y mitigar las limitaciones descritas anteriormente.

2. Metodología

2.1 Planteamiento del problema

El objetivo fundamental del trabajo realizado es la detección y localización automática de la válvula mitral durante el ciclo cardíaco. Este problema se va a abordar llevando a cabo la localización de tres puntos anatómicos de la válvula cardíaca correspondientes a las inserciones de los velos valvulares. Geométricamente, estos tres puntos, puesto que no estarán alineados, definirán el plano que contiene la válvula mitral.

En este trabajo se propone llevar a cabo la localización de dichos puntos en los planos ortogonales que definen los ejes largos del corazón (sagital y coronal) dado un volumen cardíaco reformateado según la dirección axial del corazón. Se supone por tanto necesario un reformateo previo del volumen cardíaco según el eje del ventrículo izquierdo. Este paso previo se puede realizar de forma automática en estaciones de trabajo de CT cardíaco.

Por tanto, para llevar a cabo la localización y detección de estos tres puntos que definen el plano valvular mitral, se propone la siguiente metodología compuesta de dos fases diferenciadas. La primera fase consiste en la detección de aquellas vistas bidimensionales (ejes largos del corazón, sagital y coronal), dado el volumen cardíaco en 3D reformateado, que contengan los puntos de interés. Esta tarea se llevará a cabo mediante una Red Neuronal Convolutiva (CNN) de clasificación que seleccionará el plano donde se realizará la posterior localización de los puntos de interés. La segunda etapa del método propuesto consiste en la detección de los puntos de interés, dado los planos predichos y localizados de la fase anterior. Para ello se propone el uso de una segunda CNN de detección, que detectará y localizará los puntos en los planos seleccionados por la primera CNN.

Esta operación se repetirá para cada una de las vistas (sagital y coronal). Dos de los puntos de interés se localizarán en el eje largo sagital del corazón, mientras que el tercero se detectará en eje largo coronal. En definitiva, se utilizarán dos CNNs de clasificación (una para cada vista) y tres CNNs de detección (una para cada punto). De esta forma, podremos detectar de forma automática las coordenadas tridimensionales de los 3 puntos que definen el plano valvular mitral. Esta metodología en dos fases queda definida en la Figura 1.

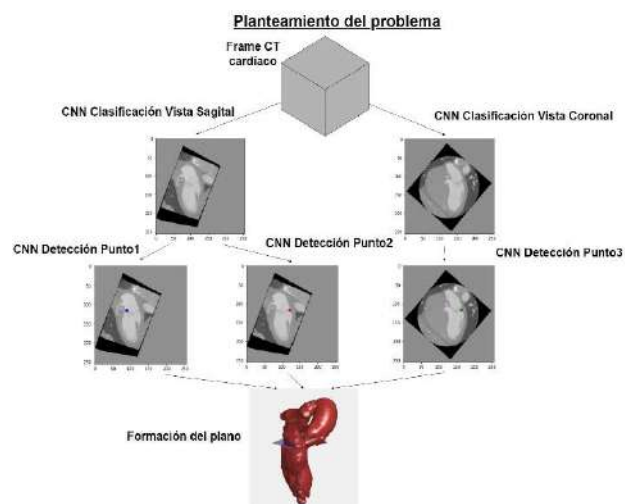


Figura 1. Representación esquemática del planteamiento del problema.

2.2 Red Neuronal Convolutiva de clasificación

Como se ha explicado anteriormente, en esta primera etapa, una primera CNN de clasificación seleccionará aquellas vistas bidimensionales en los ejes largos donde posteriormente se podrá llevar a cabo la detección de los puntos de interés. Esta red de clasificación recibirá como entrada cada una de las vistas bidimensionales que contenidas en el volumen cardíaco para poder ser clasificadas de forma binaria como vistas “aptas” o “no aptas” para la detección de los puntos. Estas imágenes bidimensionales tienen un tamaño de 256x256 píxeles. A esta capa de entrada le siguen dos capas convolucionales, con un banco de 32 filtros cada una, todos ellos de tamaño 3x3 para captar información local de la imagen. Estas primeras capas serán las encargadas de aprender las

características de bajo nivel de los datos de entrada. A estas capas convolucionales, le sigue una capa de submuestreo de tipo *max-pooling* con un campo receptivo de 2x2, seguida de otras dos capas convolucionales (64 filtros de 3x3) y otra de submuestreo. Una vez llegados a este punto, la red habrá extraído y aprendido las características directamente de los datos de entrada para su correcta clasificación. Cabe destacar que después de cada capa convolutiva se aplica a la salida la función no lineal ReLU.

La arquitectura finaliza con una sucesión de tres capas densamente conectadas (Perceptrón Multicapa) cuya función es unificar toda la información de las características extraídas de las capas convolucionales. La última capa, tendrá dos neuronas (clasificación binaria anteriormente comentada) con una activación no lineal tipo *softmax* que proporciona las probabilidades de que la entrada pertenezca a cada una de las clases.

La arquitectura de la CNN de clasificación se puede observar en la Figura 2a.

2.3 Red Neuronal Convolutiva de detección

Una vez seleccionada la vista óptima para la detección de los puntos de interés, se predecirán las coordenadas de los mismos mediante una Red Neuronal Convolutiva de detección.

En esta red de detección o regresión de las coordenadas de los puntos de interés, las capas convolucionales y de agrupamiento tienen el mismo tamaño que en la red de clasificación y en su salida se aplica también la función no lineal ReLU. Sin embargo, esta red de detección es más compleja y profunda, formada por 12 capas convolucionales y 5 de agrupamiento, ya que las características discriminativas para la detección de los puntos son más complejas. Finalmente termina con tres capas densamente conectadas cuya capa de salida está formada por dos neuronas correspondientes a las dos coordenadas predichas (ver Figura 2b). En este caso no se hará uso de la función *softmax* ya que esta CNN es una red de regresión que tiene como salida la estimación de las coordenadas x, y del punto que se desea localizar.

Cabe destacar que tanto en la red de clasificación como en la de detección, el entrenamiento se basa en un problema de optimización para la minimización de la función de coste, en clasificación se optimiza la entropía cruzada y en

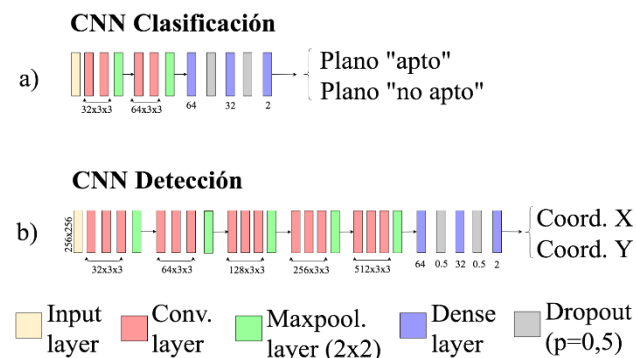


Figura 2. Configuraciones de las arquitecturas de las CNN de clasificación (a) y detección (b)

detección el error cuadrático medio. En los dos casos se utiliza el método de descenso de gradiente estocástico (SGD) con una actualización de momentos de Nesterov [2].

2.4 Prevención del sobreajuste

Las Redes Neuronales Convolucionales, que involucran miles de parámetros, tienden fácilmente al sobreajuste de los mismos a la base de datos de entrenamiento, impidiendo una correcta generalización. Este problema se ha abordado utilizando cuatro técnicas diferentes.

En primer lugar, se ha utilizado la técnica de la interrupción anticipada, la cual detiene el entrenamiento de la red antes de que empiece el sobreajuste [3].

En segundo lugar, se ha aplicado una regularización de la función de coste. Mediante la regularización L2 se penaliza la magnitud al cuadrado de todos los pesos de la red (w) directamente en la función de coste añadiendo un término regularizador de valor $\frac{1}{2}\lambda w^2$, donde λ es el parámetro regularizador [4]. Este método penaliza los cambios abruptos dando preferencia a los cambios suaves en la dirección del gradiente de error.

También se ha aplicado la técnica de *dropout*, consistente en desconectar aleatoriamente ciertas neuronas a lo largo de la arquitectura con una cierta probabilidad durante el entrenamiento [5]. Esta práctica es equivalente a entrenar un conjunto de redes que contenga todas las posibles configuraciones de dicha red. De esta forma, se disminuirá el sobreajuste.

Por último, se ha aplicado la técnica de aumento de datos. Esta técnica se basa en aumentar el conjunto de datos para que haya mayor variabilidad en el conjunto de datos de entrenamiento provocando que sea más difícil el sobreajuste. En este caso, se han multiplicado por 5 el número de imágenes de entrenamiento mediante combinaciones de transformaciones de rotación ($\pm 20^\circ$), escala ($\pm 10\%$) y traslación en ambos ejes de la imagen ($\pm 20\%$ de la longitud de los mismos).

3. Experimentos y Resultados

El conjunto de datos utilizado en este trabajo consiste en 96 *frames* (sin el aumento de datos aplicado) de CT cardiacos correspondientes a la fase de sístole provenientes de 12 pacientes diferentes (8 *frames* por paciente). Para cada uno de los pacientes y para cada *frame* de los mismos, se llevó a cabo una localización manual a través de la herramienta OsiriX de los tres puntos de interés conformando así la base de datos de entrenamiento y validación de la herramienta propuesta. Cabe destacar que se han eliminado del estudio aquellos *frames* correspondientes a la fase de diástole debido a la gran dificultad de detectar visualmente los puntos de interés causada por la apertura de la válvula durante esta fase del ciclo cardiaco.

3.1 Metodología de evaluación y justificación de hiperparámetros

Para realizar una evaluación de los experimentos y poder justificar la elección de los hiperparámetros se han dividido

los datos de los 12 pacientes en 3 grupos. Por un lado, tres pacientes forman el conjunto de *test* para la evaluación de la metodología, mientras que 9 pacientes se reservaron para el entrenamiento y la elección de hiperparámetros de las mismas (conjunto de entrenamiento y validación). La separación de pacientes para el entrenamiento y validación se realizó utilizando el método de validación cruzada de 9 iteraciones, asegurando que 8 pacientes se utilizan para training y 1 para validación en cada iteración.

Mediante esta metodología se pudo optimizar las arquitecturas de ambas redes de clasificación y detección, pudiendo llegar a las presentadas en la secciones 2.2 y 2.3, comprobando también un mejor comportamiento cuando se hace uso de redes de detección individualizadas para cada punto en lugar de una misma red de detección para los 2 puntos de la vista sagital.

3.2 Evaluación en el conjunto de *test*

El conjunto independiente de datos de *test* se compone de 24 *frames* provenientes de 3 pacientes independientes a los del entrenamiento. Tras predecir los puntos de interés mediante la metodología en dos fases anteriormente expuesta, se evaluará la precisión en la detección de los mismos para cada paciente y punto bajo estudio, reportando la distancia media (mm) entre los puntos predichos y reales. Adicionalmente, se reportará también la distancia según el eje largo del corazón (Error(z)) entre los planos predichos y los reales, ya que la distancia entre los planos en esta componente es la que afectará en mayor medida al cómputo de parámetros derivados funcionales.

Paciente	Error [mm]	Error (z) [mm]	Punto	Error [mm]
#1	7,26	1,52	#1 (sagital)	6,41
#2	16,39	11,43	#2 (sagital)	7,22
#3	5,46	1,48	#3 (coronal)	5,43
Media	9,70	4,81	Media	9,70
(std)	(5,17)	(4,98)	(std)	(5,17)

Tabla 1. Resultados de la evaluación del conjunto de *test*.
Izquierda: precisión por paciente bajo estudio. Derecha: precisión por punto bajo estudio.

Como se desprende en los resultados presentados en la Tabla 1, se puede observar cierta variabilidad en el rendimiento del esquema de procesamiento presentado en función de los distintos pacientes. En los pacientes primero y tercero, se obtiene un error en la componente transversal inferior a 2mm. Sin embargo, se comete un error de 11,43mm en el segundo caso. Esto es debido a que este paciente en concreto presentaba una anatomía compleja y diferente en relación a los datos de entrenamiento, siendo por tanto más compleja la detección de los puntos. Por otro lado, si se analiza la detección de cada punto por separado, se observa que el mayor error se obtiene durante la localización de los puntos correspondientes a la vista sagital (puntos 1 y 2), aspecto que se tendrá en cuenta de cara a futuras mejoras, ya que un error en localización de los puntos conllevará un fallo en la localización del plano valvular mitral.

Por otro lado, cabe destacar que medidas cuantitativas de error de la distancia global entre los puntos predichos y etiquetados manualmente pueden no corresponder con la medida de precisión en la localización del plano valvular, ya que un mismo plano puede estar definido por cualquiera de los infinitos puntos que lo contienen. Para un mayor análisis, y aparte del cálculo de la distancia de los planos según el eje largo del corazón (Error(z) en la Tabla 1), se ha llevado a cabo una evaluación visual de los resultados obtenidos por la herramienta propuesta. Estos resultados se muestran en la Figura 3, donde se muestran 4 *frames* de los 3 pacientes de test incluyendo la segmentación y reconstrucción tridimensional del ventrículo izquierdo, aurícula izquierda y aorta de los *frames*, junto con los planos localizados manualmente (azul) y el estimado por la herramienta (verde).

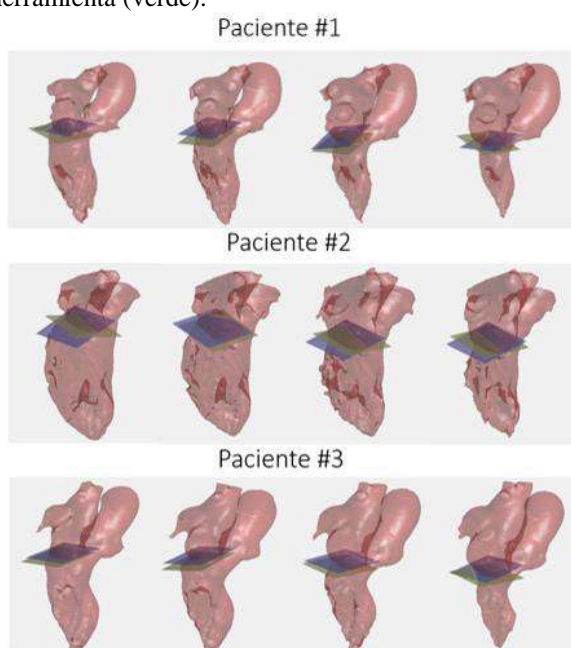


Figura 3. Resultados visuales de la predicción del plano mitral. Azul: Plano definido manualmente, verde: plano estimado a través de la herramienta propuesta.

En este caso, los resultados cuantitativos de error concuerdan con los resultados visuales, ya que los pacientes #1 y #3 con menor error son los que, visualmente, tienen una mejor detección del plano. Cabe destacar también que en aquellos casos donde falla la detección de un solo punto, la orientación del plano se convierte en errónea. Esto sobre todo ocurre en la localización de los puntos del plano sagital, donde normalmente se comete un mayor error, provocando un desplazamiento indeseado en la detección del plano. Los resultados fallidos obtenidos en el paciente de test #2 se pueden asumir a la diferente morfología y anatomía del corazón respecto a los datos de entrenamiento. Queda por tanto también destacada la necesidad de llevar a cabo un entrenamiento de la herramienta propuesta con un número mayor de datos, para poder recoger la mayor variabilidad anatómica posible. Aun así, los resultados son prometedores debido a que en la mayoría de los casos localiza la posición del plano correctamente con un error inferior a 2mm.

4. Conclusiones

Este trabajo propone un método basado en Redes Neuronales Convolucionales (CNN) para la detección automática del plano valvular mitral a partir de imágenes de CT cardíaco dinámico. Para abordar este problema, se propone localizar de forma automática tres de las intersecciones de los velos de la válvula mitral, que a su vez definirán el plano valvular. En cuanto a la resolución de esta localización del plano, se ha abordado el problema en dos fases diferenciadas, por un lado, la detección de aquellos planos óptimos para llevar a cabo la detección de los puntos de interés, y por otro, dado estos planos, la detección en si misma de los puntos. Para ello se han implementado y definido dos Redes Neuronales Convolucionales (CNN) específicas para la clasificación de las vistas óptimas y regresión de las coordenadas de los puntos bajo estudio.

Ambas CNN propuestas han sido entrenadas, optimizadas y validadas mediante una técnica de validación cruzada de 9 iteraciones, utilizando 72 *frames* provenientes de 9 pacientes diferentes. Por otro lado, la herramienta propuesta ha sido evaluada en un conjunto de *test* independiente en 24 *frames* de 3 pacientes diferentes, obteniendo resultados prometedores en la tarea de la localización automática del plano valvular mitral dado una secuencia dinámica de CT cardíaco. Para mitigar los posibles errores de la herramienta presentada, se propone también re-entrenar los algoritmos propuestos con un mayor número de imágenes para poder así contar con mayor variabilidad en la base de datos de entrenamiento y poder minimizar el error en la localización del plano valvular mitral en aquellos casos con anatomía compleja. También se propone el uso conjunto de CNN y Redes Neuronales Recurrentes (RNN) para incorporar conocimiento espacio-temporal en la localización del plano y para la extensión de la misma a los *frames* de diástole a partir de la detección en los *frames* de sístole. Las dependencias espaciales quedarán modeladas por las CNN, mientras que las dependencias temporales de la posición de los puntos bajo estudio serán capturadas por las RNN, pudiendo llegar así a resultados más precisos.

Referencias

- [1] V. Bertomeo, et al. Situación de la enfermedad cardiovascular en España. Del riesgo a la enfermedad. *Revista Española de Cardiología*, Vol. 8, Núm. Supl. E, pp 2-9, 2008.
- [2] Y. Bengia, et al. Advances in optimizing recurrent network. *International Conference on Acoustics, Speech and Signal Processing (IEEE)*, pp 8624-8628, 2013.
- [3] Y. Yao, et al. On early stopping in gradient descent learning. *Constructive Approximation*, Vol. 26, pp 285-312, 2005.
- [4] C. Cortes, et al. L2 Regularization for Learning Kernels. *Conference on Uncertainty in Artificial Intelligence (UAI)*, Montreal, Quebec, 2009.
- [5] N. S. S. S. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning*, pp, 1929-1958, 2011.