VII. Conclusiones y líneas futuras de investigación

La evaluación sistemática del desempeño de algoritmos de segmentación es un apartado esencial de la validación de los algoritmos que se desarrollan en el ámbito de la segmentación para el que han sido diseñados. Sin embargo, no se deposita tanto esfuerzo en la investigación de mejoras para la evaluación como en el desarrollo de nuevos algoritmos, como se ha expuesto en la sección dedicada a las Métricas de evaluación para algoritmos de segmentación. De hecho, no se ha resuelto el problema del todo, a pesar de haberse propuesto marcos unificados de evaluación, como se estudió en la citada sección.

La evaluación realizada en este proyecto no pretende ser generalista, ni pueden extraerse de ella conclusiones con significancia estadística. No obstante, aclara qué algoritmos, de entre algunos que ofrece la plataforma ITK, son más apropiados para resolver el problema la segmentación de la aurícula izquierda a partir de imágenes de TAC en entornos hospitalarios. Como se analiza en profundidad en la sección Comparativa de los algoritmos estudiados, los algoritmos más apropiados de entre los estudiados siguiendo la metodología descrita, como primera opción para la segmentación de la aurícula izquierda en entornos hospitalarios, son los algoritmos de crecimiento de región, en particular los algoritmos ConfidenceConnected y IsolatedConnected. ThresholdLevelSet consigue resultados muy cercanos a estos dos primeros, y en algún caso mejores, perdiendo en cambio eficiencia y facilidad en el ajuste de parámetros de una manera significativa. OtsuMultipleThreshold se ha mostrado como un algoritmo no apropiado para la tarea de la segmentación de la aurícula izquierda, consiguiendo resultados bastante deficientes.

A continuación se discuten diversas líneas abiertas para la investigación que el presente trabajo no ha desarrollado.

- Sería interesante extender la evaluación a todos los algoritmos de segmentación que ofrece ITK, para disponer de una valoración completa sobre la plataforma en su aplicación a la segmentación de la aurícula izquierda en la práctica clínica.
- Sería útil extender la base de imágenes TAC disponible para que los resultados obtenidos en la evaluación poseyeran significatividad estadística.
- Asimismo, sería interesante extender la evaluación a otras técnicas de imagen, como la resonancia magnética.
- Se podría también pensar en realizar la evaluación extendida al caso 3D, es decir, midiendo la eficacia de los algoritmos al segmentar los volúmenes completos de la aurícula izquierda en lugar de solamente cortes axiales de la misma.
- Otra posible mejora consistiría en introducir en la evaluación el efecto de la diferencia introducida por los expertos al realizar las segmentaciones *ground truth*. Esto podría llevarse a cabo implementando el método *STAPLE* (*Simultaneous truth and performance level estimation*) (30).
- Durante la observación de los resultados obtenidos en este proyecto, se ha puesto de manifiesto una mejora para los algoritmos de crecimiento de regiones evaluados. Durante el crecimiento de región, es frecuente que existan desviaciones de contraste más grandes de las consideradas por el algoritmo aún dentro de la aurícula, de manera muy localizada y en un área que ocupa un número escaso de píxeles. Este efecto resulta en una subsegmentación de la aurícula, estando el área segmentada formada por pequeños agujeros que no se consideran parte de la aurícula cuando sí lo son. Podría emplearse el conocimiento de que la aurícula es una estructura conexa, excepto por las venas pulmonares y la orejuela, según el corte axial que se considere. Pero en el caso de estas estructuras, el tamaño de las mismas es característico y siempre superior a los agujeros observados. Así, sería interesante evaluar la mejora que se introduciría en los resultados de la segmentación en

los algoritmos de crecimiento de región tras aplicar una función que rellenara los agujeros menores que una cantidad específica característica que también habría que ensayar empíricamente.