

Segmentación automática de grasa parafaríngea en imágenes de resonancia magnética usando U-net 3D

Oscar Ivan Duarte Torres^a, Daniel Sebastián Chaves Burbano^b, Liliana Otero Mendoza^c, Mariana Saltos Cristiano^d, Luis Felipe Uriza Carrasco^b, María Daniela Alvarado Lasso^c, Cesar Julio Bustacara Medina^a

^aFacultad de Ingeniería, Pontificia Universidad Javeriana, Bogotá, Colombia

^bFacultad de Medicina, Pontificia Universidad Javeriana, Bogotá, Colombia

^cFacultad de Odontología, Pontificia Universidad Javeriana, Bogotá, Colombia

^dHospital Universitario San Ignacio, Bogotá, Colombia

Abstract

Justificación y Objetivo: La segmentación de estructuras y tejidos en las vías aéreas superiores provee herramientas para encontrar las posibles causas de la Apnea Obstructiva del Sueño (AOS) y su relación con dichas estructuras. El tejido graso en la región parafaríngea, puede tener una relación con la obstrucción de la vía aérea en pacientes con AOS. Diseñar una red neuronal convolucional para la detección y segmentación del tejido graso en la región parafaríngea.

Materiales y métodos: En el estudio se utilizó un dataset de 64 imágenes de resonancia magnética ponderadas T1, seleccionadas de forma aleatoria de un total de 140 imágenes, y sus respectivas segmentaciones de grasa parafaríngea, realizadas de forma manual. A partir de la selección de las imágenes se realizó el preprocesamiento de éstas mediante recorte y re-espaciado principalmente. Se realizó el entrenamiento de la red neuronal, se segmentaron las imágenes y la segmentación se evaluó mediante coeficiente Dice. Finalmente se calculó el volumen de la grasa parafaríngea y se comparó con el volumen de la segmentación manual.

Resultados: Se obtuvo un coeficiente Dice promedio de 0.8086 para la red U-net 3D con función de activación ReLU y de 0.6367 con la misma red, pero usando la función de activación sigmoide.

Conclusiones: La red neuronal Unet 3D implementada permite la segmentación automática de la grasa parafaríngea y tiene un potencial uso en encontrar relaciones entre la presencia de grasa parafaríngea y la existencia y

severidad de la AOS, presentando resultados superiores que en otros estudios en cuanto a coeficiente Dice.

Keywords: Segmentación de imágenes, Procesamiento de imágenes médicas, Redes convolucionales, Aprendizaje profundo

INTRODUCCIÓN

Las vías aéreas superiores comprenden una parte del sistema respiratorio que va desde los labios y las fosas nasales hasta la tráquea. En el contexto de los estudios llevados a cabo mediante el uso de imágenes médicas en las vías aéreas, se encuentran estudios relacionados con trastornos del sueño como la apnea obstructiva del sueño. La apnea obstructiva del sueño (AOS), es un trastorno respiratorio del sueño, en el cual ocurre estrechamiento parcial o completo, o colapso, de las vías respiratorias faríngeas durante el sueño, que produce una interrupción repetida del flujo de aire, desaturación de oxígeno e interrupción del sueño [1].

La apnea obstructiva del sueño es un problema importante de salud pública [2] porque es un factor de riesgo para enfermedades cardiovasculares como la hipertensión arterial, falla cardíaca, arritmia y enfermedad coronaria [3]. La prevalencia de AOS es alta y se puede incrementar debido a la epidemia actual de obesidad. La AOS tiene efectos económicos, aumenta los accidentes de tránsito, y disminuye la calidad del trabajo [4].

Los estudios mediante métodos no invasivos, como las imágenes médicas se encaminan a realizar la detección de obstrucciones y caracterización de las vías aéreas superiores con métodos automáticos o semiautomáticos de segmentación de imágenes, para apoyar los procesos de diagnóstico y/o tratamiento de la Apnea Obstructiva del Sueño [5, 6, 7, 8]. Estos estudios se realizan usando imágenes de Tomografías Computarizadas, Resonancia Magnética, ultrasonido y recientemente con Tomografías Computacionales de haz cónico o CBCT (Cone Beam Computed Tomography) por su sigla en inglés.

Hay varios estudios que sugieren la existencia de una relación entre la AOS y los tejidos grasos en las vías aéreas superiores, en los cuales se investiga si la acumulación de grasa en los tejidos que rodean la vía área superior retropalatina contribuyen a la aparición de la AOS. Estos estudios fueron realizados con datos obtenidos a partir de medidas antropométricas [1, 9], y de imágenes de resonancia magnética [8, 10].

A partir de la relación aparente mencionada anteriormente, ha surgido el interés en el estudio de la acumulación de grasa en la lengua y en la región retropalatina de pacientes con AOS, encontrando que la grasa depositada en los tejidos que rodean la vía aérea superior contribuyen a la colapsabilidad de la vía aérea en personas con y sin AOS [8, 11, 12, 13].

Los estudios de las vías aéreas superiores a nivel general presentan usos variados, orientados principalmente en encontrar relaciones entre la anatomía de las diferentes estructuras en la región como la lengua, la faringe, el volumen de los tejidos grasos, entre otros, y la existencia de la AOS. Por esta razón surge el interés en investigar si es posible establecer la existencia de una relación entre la acumulación de grasa en la región parafaríngea y en la lengua con la severidad de la AOS en pacientes adultos.

La búsqueda de la respuesta a la pregunta anterior se encuentra bajo investigación por parte de un equipo interdisciplinario de odontólogos, radiólogos e ingenieros de la Pontificia Universidad Javeriana y el Hospital Universitario San Ignacio. En la búsqueda de aportar a la investigación, en el presente estudio se planteó la siguiente pregunta: ¿Cómo detectar y segmentar de forma automática el tejido graso en la región parafaríngea a partir de imágenes de Resonancia Magnética? Con el método de segmentación automática implementado, se puede ayudar en investigaciones futuras a identificar la relación entre la acumulación de grasa en la región parafaríngea con las obstrucciones de la vía aérea en individuos con AOS.

MATERIALES Y MÉTODOS

A partir de las fases mencionadas anteriormente, en esta sección se describe el proceso que se llevó a cabo para realizar la segmentación de grasa parafaríngea, que se puede observar en la Figura 1 y los artefactos producto de cada tarea en el proceso.

Caracterización y selección de imágenes

El objetivo de esta fase fue extraer información de las imágenes relacionada con el tamaño de la imagen, tamaño de los voxels o spacing e intensidades. El producto que se obtuvo en esta fase es una tabla consolidando mínimos, máximos y media de cada uno de los ítems mencionados anteriormente. A partir de la información de las imágenes se seleccionó de forma aleatoria la muestra de 64 imágenes que hacen parte del estudio.

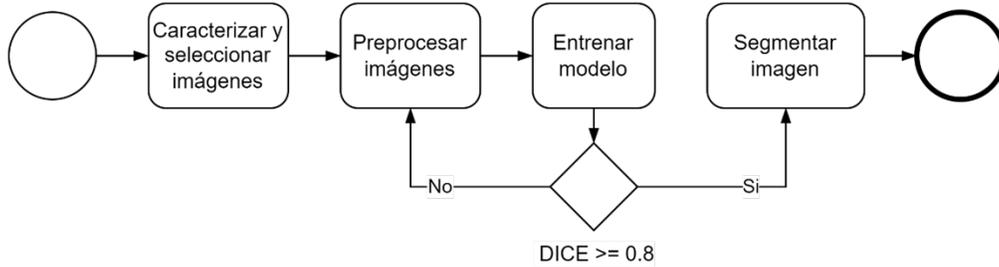


Figura 1: Proceso de segmentación de grasa parafaríngea

Caracterización de las imágenes

Se obtuvieron 142 imágenes de resonancia magnética de las cuales se excluyeron 2 por no tener la calidad suficiente para el estudio debido a artefactos o baja resolución. A partir de las restantes 140 se realizó una caracterización de las imágenes obteniendo los valores mínimos, máximos y medias para los tamaños en los ejes x, y, z, el spacing en cada uno de los ejes y las intensidades mínimas y máximas encontradas en las imágenes de resonancia de magnética.

Tabla 1: Caracterización de imágenes consolidada

	x	y	z	Sx	Sy	Sz	IMin	I _{max}
Media	341	484	484	0.50	0.50	0.50	0	13980
Mínimo	320	480	480	0.50	0.48	0.48	0	781
Máximo	360	528	528	0.50	0.50	0.50	0	65533

En la Tabla 1 se observan los valores anteriormente descritos de la siguiente manera:

- X: Cantidad de voxeles en el eje x
- Y: Cantidad de voxeles en el eje y
- Z: Cantidad de voxeles en el eje z
- Sx: Spacing eje x
- Sy: Spacing eje y
- Sz: Spacing eje z

- Imin: Intensidad mínima
- Imax: Intensidad máxima

Selección de las imágenes

Para realizar la selección de las imágenes se tuvo en cuenta el grupo al que pertenecía el sujeto de la imagen que puede ser paciente o control. Los pacientes son el grupo de personas que han sido diagnosticados con apnea, mientras que los sujetos pertenecientes al grupo de control, no. Cada sujeto del estudio tiene 4 imágenes de resonancia: resonancia T1 de alta resolución, resonancia T1 de baja resolución, resonancia T2 y resonancia T2 FLAIR. Las imágenes seleccionadas para esta investigación, cuya caracterización fue presentada anteriormente, son las T1 de alta resolución ya que en estas el tejido graso presenta altas intensidades y lo hacen mayormente visible y permiten obtener más información sobre la imagen que las de baja resolución. La muestra de 64 imágenes utilizadas en la investigación se escogió de forma aleatoria teniendo en cuenta que 32 fueran del grupo de control y 32 del grupo de pacientes. Las imágenes de la muestra descrita fueron segmentadas de forma manual utilizando el software ITK Snap por parte de un grupo de radiólogos.

Preprocesamiento de imágenes

Se prepararon las imágenes de la muestra y sus segmentaciones manuales previamente verificadas por un radiólogo para ser utilizadas en el entrenamiento y validación del modelo, recortando las imágenes para que tengan un tamaño uniforme y normalizando las intensidades. El producto de esta tarea es el conjunto de imágenes y segmentaciones preprocesadas. Para realizar el entrenamiento y las segmentaciones se implementó el flujo de preprocesamiento que se puede observar en la Figura 2.

La transformación de Orientación RAS+ permite cambiar la orientación de las imágenes a una orientación estándar RAS(Right-Anterior-Superior) del eje de coordenadas desde la orientación LPI(Left-Posterior-Inferior) que tenían las imágenes. El recorte de la imagen a un tamaño de 224x384x352 voxeles se realizó para reducir la imagen a una región de interés donde se encuentran los fat pads. El cambio del tamaño o espacio de los voxeles a 1x1x1 se realizó teniendo en cuenta que todas las imágenes tienen un espacio de voxeles de 0.5x0.5x0.5 o muy cercanos a este valor para reducir el tamaño de las imágenes a la mitad. El último recorte a un tamaño de 96x192x160 se

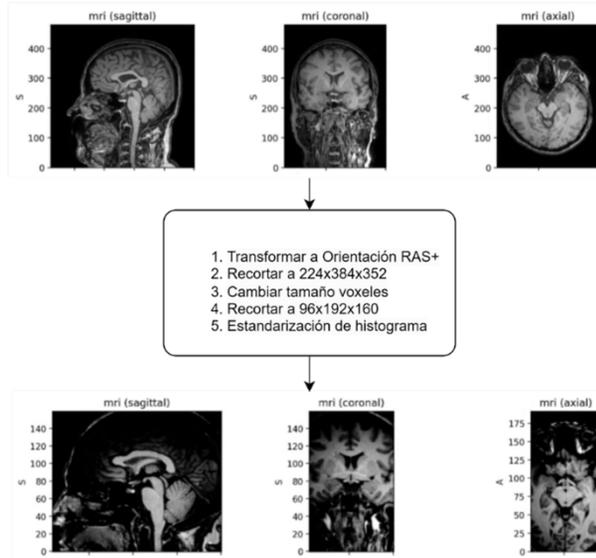


Figura 2: Transformaciones preprocesamiento

realizó para ajustar el tamaño a múltiplos de 32 para lograr compatibilidad con la red neuronal evitando afectar al máximo la región de interés donde se ubican los fat pads. Para las segmentaciones de entrenamiento y validación se aplicó la transformación de codificación oneHot, para evitar que el sistema tome el valor de las etiquetas (0: fondo, 1: grasa parafaríngea) como valores numéricos, haciéndolos categóricos.

La última transformación de aplicada a las imágenes fue la estandarización del histograma, debido a la dispersión de la intensidad máxima identificada en la caracterización de las imágenes y que se puede observar en la Figura 3. A partir de la estandarización se obtienen valores de intensidad con una mayor uniformidad para cada una de las imágenes. Las transformaciones aplicadas descritas en esta sección son parte de la implementación propuesta por [14].

Entrenamiento del modelo

Se utilizaron las imágenes preprocesadas y los hiperparámetros de la configuración del modelo, se entrena el modelo teniendo como función de optimización la minimización de la pérdida de coeficiente de DICE que implica la maximización del coeficiente. Si en la validación del modelo se obtiene un

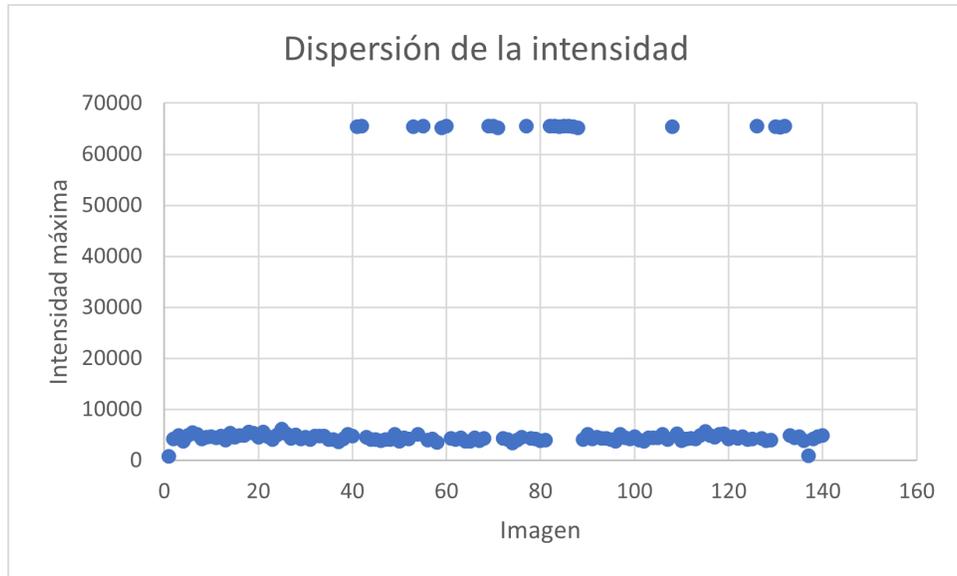


Figura 3: Gráfica de dispersión de intensidades máximas.

coeficiente DICE medio mayor o igual a 0.8 se considera finalizado el entrenamiento. En caso de que no se obtenga el coeficiente Dice esperado se procede a realizar un nuevo preprocesamiento de imágenes y modificación de hiperparámetros. El producto resultante de esta tarea son los pesos de la red neuronal que se utilizan en el modelo para la segmentación de nuevas imágenes. Arquitectura de la red neuronal La arquitectura de la red U-net 3D diseñada en este proyecto se encuentra basada en la descrita en [15], pero por detalles de implementación se realizó una reducción de los canales de salida de cada capa pasando de 32 a 8 canales salida de la primera capa y utilizando la función de pérdida de coeficiente DICE la cual viene dada por $1 - \text{DICE}$, buscando minimizar la pérdida para obtener mejores resultados.

En la Figura 4 se puede observar el diagrama de la arquitectura de la red neuronal implementada, en esta arquitectura a parte de la diferencia entre los canales de salida también se maneja un número de clases distinto, en este caso se manejan 2 clases, el fondo y los fat pads. La imagen de entrada a la red neuronal es del tamaño de 96x192x160 voxels del mismo modo que la de salida. Para la implementación de la arquitectura diseñada se utilizó la librería unet, que permite la implementación de redes U-net en dos y tres dimensiones, propuesta por [16].

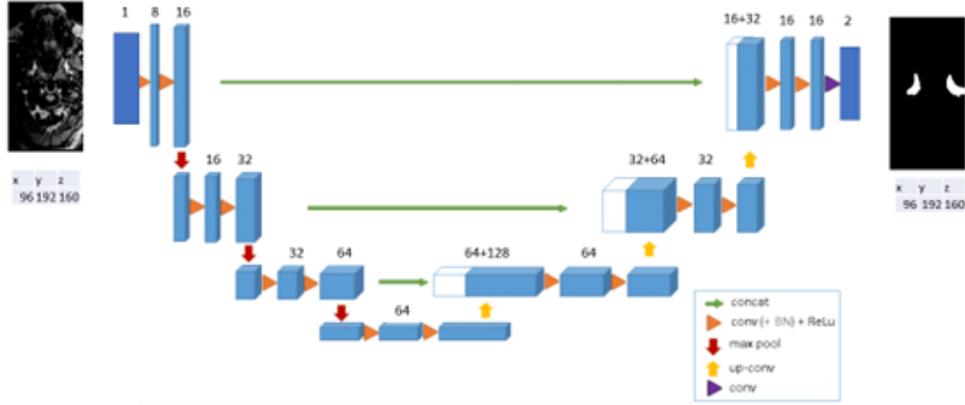


Figura 4: Arquitectura U-net 3D implementada. Adaptada de [15]

Segmentación de la imagen

el modelo en funcionamiento recibe una imagen y devuelve la segmentación de la grasa para faríngea en dicha imagen. Coeficiente de Similitud Dice El coeficiente de similitud Dice (DICE), también llamado el índice de superposición, es la métrica más usada para la validación de segmentación de volúmenes médicos [17]. El coeficiente DICE mide la superposición espacial entre dos segmentaciones, definido como $DSC(A, B) = 2(A \cap B)/(A + B)$ donde \cap es la intersección y A y B son regiones [18]. La fórmula para calcular el índice DICE se define así:

$$DICE = \frac{2(A \cap B)}{A + B} = \frac{2TP}{2TP + FP + FN}$$

Siendo TP los verdaderos positivos, FP los falsos positivos y FN los falsos negativos. A partir de la fórmula anterior se obtiene un resultado entre 0 y 1, siendo 0 una superposición nula y 1 una total superposición entre las dos segmentaciones.

RESULTADOS

Para la realización del entrenamiento y segmentación se utilizó un equipo con procesador Intel Xeon W-10885M CPU de 8 núcleos de 2.40GHz, con 32 GB de RAM. Debido al límite de memoria se aplicaron las transformaciones, para la reducción del tamaño de las imágenes, durante la etapa de

preprocesamiento descrita anteriormente. Para el entrenamiento del modelo se utilizaron 44 imágenes, que corresponden al 70 % de la muestra y para la validación se utilizaron 20 imágenes, que corresponden al 30 % de la muestra. Para la prueba con la cual se evaluó la segmentación se utilizó el 100 % de la muestra. En total se realizaron 2 pruebas una entrenando un modelo con función de activación ReLu y otra con función de activación Sigmoide.

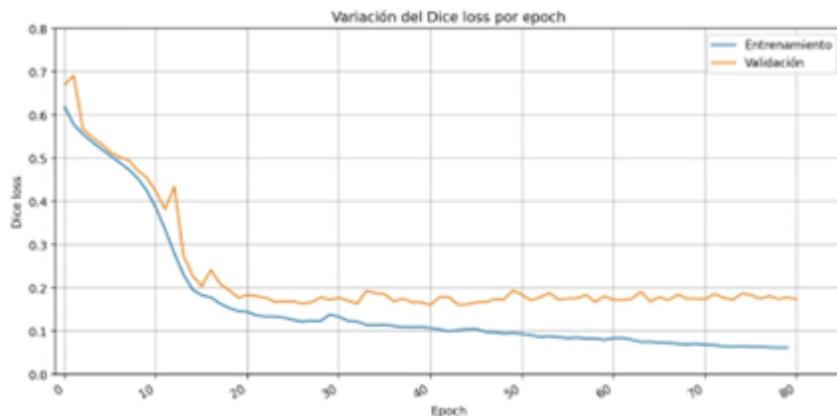


Figura 5: Perdida de DICE en el entrenamiento y validación con ReLu

En la Figura 5 se observa la variación de la métrica de perdida de Dice utilizada para el entrenamiento y validación del modelo con arquitectura U-net 3D con la función de activación ReLu, la cuál es la función usada en la definición original de dicha arquitectura.

En la Figura 6 se puede observar la variación de la métrica de perdida de Dice utilizada para el entrenamiento y validación del modelo con arquitectura U-net 3D con la función de activación Sigmoide. La función sigmoide se utilizó para verificar si se podía obtener una mejoría en la segmentación con respecto a ReLu debido a las características morfológicas de la región segmentada.

Tabla 2: Comparación de métricas para la validación, entrenamiento y prueba

	Dice Loss Tr (SD)	Dice Loss Val (SD)	DICE (SD)
ReLu	0.0624 (0.0076)	0.1786 (0.0172)	0.8086 (0.1152)
Sigmoide	0.1112 (0.0120)	0.2038 (0.0120)	0.6347 (0.1433)

En la Tabla 2 se observa la consolidación de las métricas obtenidas en el entrenamiento (Dice Loss Tr) y validación (Dice Loss Val), la perdida de Dice

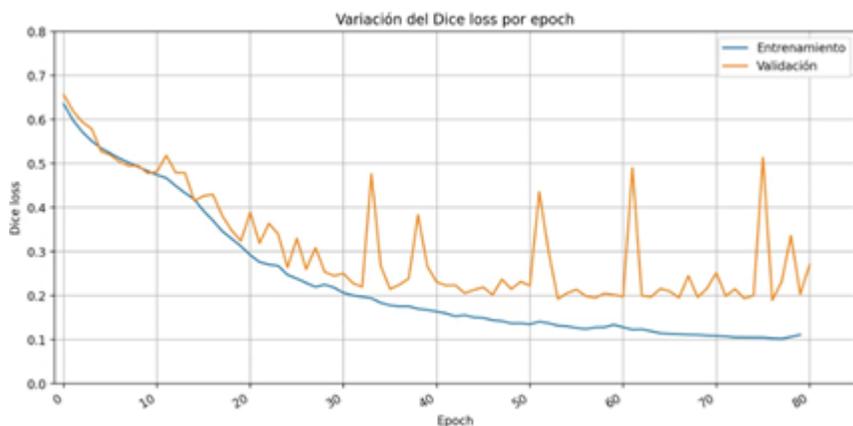


Figura 6: Perdida de DICE en el entrenamiento y validación con Sigmoide

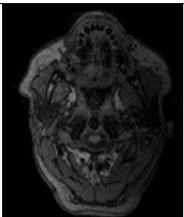
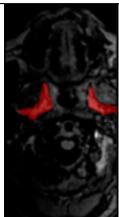
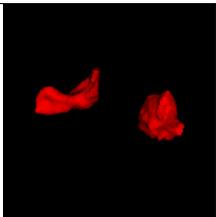
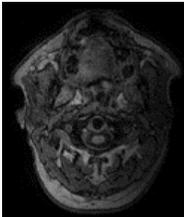
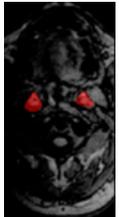
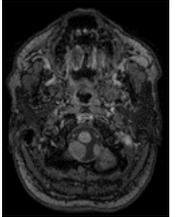
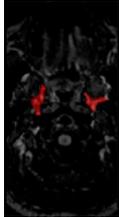
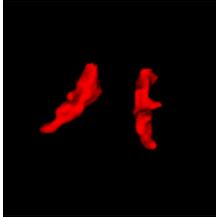
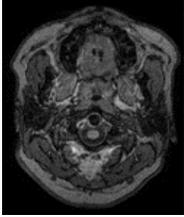
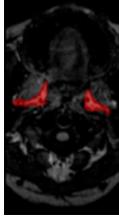
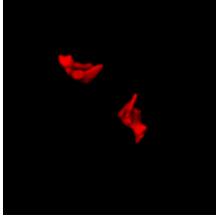
con su desviación estándar (SD) y la media de coeficiente DICE obtenido en las pruebas con su SD.

Tabla 3: Comparación volúmenes calculados

Grupo/Segmentación	Manual	Unet-3D
Promedio Pacientes	13812.7003	12577.3750
Promedio Control	10383.6875	10629.0313
Promedio General	12098.1939	11603.2031

En la Tabla 3 se puede observar el volumen promedio de la grasa parafaríngea calculado para cada uno de los grupos a los que pertenecen las imágenes, paciente y control, y el promedio de todas las imágenes. Estos cálculos se realizaron a partir de las segmentaciones obtenidas en las pruebas usando la función de activación ReLu y a partir de la segmentación manual verificada por el grupo de radiólogos.

Tabla 4: Imágenes segmentadas y volumen

Imagen Original	Imagen segmentada con U-net 3D	Volumen segmentado con U-net 3D	Volumen Sementación mm3
			15375
			11608
			9292
			10321

En la Tabla 4 se muestran algunas imágenes usadas en el estudio, la segmentación sobrepuesta sobre la imagen preprocesada, la visualización del volumen segmentado y el volumen calculado. Estas imágenes fueron tomadas a partir de la visualización de estas usando ITK Snap.

DISCUSIÓN

En la presente sección se realizará un análisis de los resultados descritos anteriormente y su comparación con estudios de segmentación de grasa para-

faríngea descritos en la revisión de literatura. Como se puede observar en la presentación de los resultados, el comportamiento de la métrica de pérdida de Dice durante el entrenamiento y validación de la red neuronal U-net 3D implementada (Figura 5 y Figura 6) fue similar para la función de activación ReLu y sigmoide, pero con mejores resultados para la primera, lo cual se ve al obtener los coeficientes Dice en la prueba realizada, con 0.8086 para ReLU y 0.6063 para la sigmoide. A partir de las segmentaciones obtenidas y teniendo en cuenta el coeficiente DICE con mayor valor, se realizó una comparación del volumen de las segmentaciones manuales y la segmentación con U-net 3D con activación ReLU.

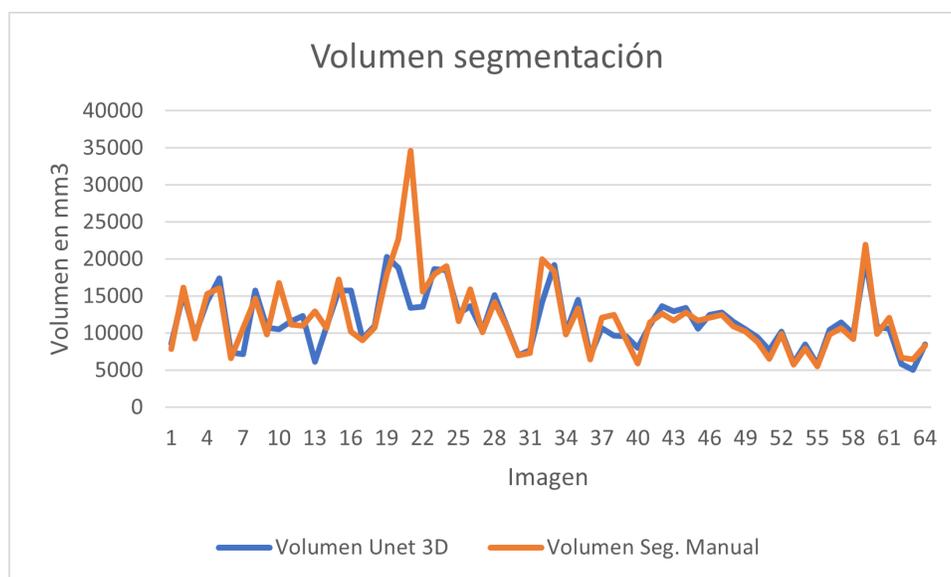


Figura 7: Volumen segmentación de cada imagen

En la Figura 7 se puede observar volumen calculado para cada una de las segmentaciones manuales y automáticas, observando una similitud. A partir de los datos de los volúmenes se calculó un coeficiente de correlación entre los dos conjuntos de datos de 0.75, lo cual indica una correlación entre el volumen de la segmentación manual y automática, pero sin significancia estadística comprobada. En cuanto a la comparación con estudios similares se tuvo en cuenta el estudio realizado por [8], que utiliza un método automático basado en análisis de componentes para segmentar la grasa para faríngea y el estudio hecho por [19], en que utilizan una red neuronal U-net modificada que usa

elementos de Resnet para segmentar diferentes estructuras en la vía aérea superior entre los que se encuentra la grasa parafaríngea.

En la Tabla 5 se puede observar el coeficiente Dice obtenido por los métodos utilizados en el presente estudio y en los estudios mencionados.

Tabla 5: Comparación resultados estudios

Estudio	Coeficiente DICE (SD)
U-net 3D ReLU (implementación propuesta)	0.809 (0.115)
Estudio en [8]	0.779 (0.041)
U-net + resnet [19]	0.580 (0.120)

El mayor coeficiente DICE se obtuvo en el presente estudio, pero con una desviación estándar mayor a la obtenida por [8], lo cual permite interpretar que el método usado en este estudio es menos exacto al del otro estudio. Comparado con el estudio de [19] se puede interpretar que el método propuesto usando U-net 3D obtuvo un mejor resultado en cuanto al coeficiente DICE y una desviación estándar similar. Cabe resaltar que los estudios comparados utilizaron distintos conjuntos de datos, número de muestras y métodos de preprocesamiento o no usaron como en [19].

CONCLUSIONES

La red neuronal Unet 3D implementada permite segmentar de forma automática la grasa parafaríngea, y se podrá utilizar en futuras investigaciones relacionadas con la AOS.

El uso de la red neuronal Unet 3D se encuentra restringido o condicionado para el entrenamiento por las capacidades de la máquina en la que se realiza su ejecución, por tal motivo, el preprocesamiento de las imágenes estuvo enfocado en la disminución del tamaño de las imágenes mediante recorte y re-espaciado.

La caracterización de las imágenes permitió identificar la diferencia en las intensidades de los voxels, lo cual conllevó a la estandarización del histograma. La red Unet 3D implementada presenta un coeficiente Dice de 0.809, el cuál es superior al de otros estudios de segmentación automática de grasa parafaríngea que presentan un coeficiente Dice de 0.779 y 0.580.

Con respecto a los trabajos futuros, para encontrar relaciones entre la presencia de grasa parafaríngea y la existencia y severidad de la AOS, se recomienda ampliar el tamaño de la muestra de pacientes con y sin AOS,

incluyendo pacientes con AOS agrupados por severidad según el IAH (índice Apnea hipopnea) para establecer la asociación entre depósitos de grasa y obstrucciones de la vía aérea en pacientes con AOS leve, moderada y severa; y pacientes con AOS agrupados por severidad según la saturación de oxígeno (T90). Para analizar estas correlaciones se recomienda utilizar técnicas de machine learning que permitan hacer múltiples comparaciones.

También se propone usar varios métodos de segmentación con el mismo dataset para poder hacer comparaciones con significancia estadística y finalmente se propone adaptar el flujo de segmentación propuesto para realizar la segmentación de grasa en la lengua que permita encontrar relaciones entre el volumen de esta estructura y la existencia de la AOS, utilizando un dataset del cual se conozcan datos de los sujetos de estudio como edad y sexo.

Referencias

- [1] R. Carter, D. Watenpaugh, Obesity and obstructive sleep apnea: Or is it osa and obesity?, *Pathophysiology* 15 (2008) 71–77. doi:10.1016/j.pathophys.2008.04.009.
- [2] K. Motamedi, A. McClary, R. Amedee, Obstructive sleep apnea: a growing problem, *Ochsner J* 9 (2009) 149–153.
- [3] M. Kato, T. Adachi, Y. Koshino, V. Somers, Obstructive sleep apnea and cardiovascular disease, *Circ J* 73 (2009) 1363–1370. doi:10.1253/circj.cj-09-0364.
- [4] N. Morsy, N. Farrag, N. Zaki, Obstructive sleep apnea: personal, societal, public health, and legal implications, *Rev Environ Health* 34 (2019) 153–169. doi:10.1515/reveh-2018-0068.
- [5] T. Ivanovska, J. Dober, R. Laqua, K. Hegenscheid, H. Völzke, Pharynx segmentation from mri data for analysis of sleep related disorders, in: G. Bebis, R. Boyle, B. Parvin, D. Koracin, B. Li, F. Porikli, V. Zordan, J. Klosowski, S. Coquillart, X. Luo, M. Chen, D. Gotz (Eds.), *Advances in Visual Computing*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, pp. 20–29.
- [6] A. Javed, Y. Kim, M. Khoo, Dynamic 3-d mr visualization and detection of upper airway obstruction during sleep using region-

- growing segmentation, *IEEE Trans Biomed Eng* 63 (2016) 431–437. doi:10.1109/TBME.2015.2462750.
- [7] M. U. R. S, C. T, I. T, Automatic pharynx segmentation from mri data for obstructive sleep apnea analysis, in: *VISAPP 2015 - 10th International Conference on Computer Vision Theory and Applications; VISIGRAPP, Proceedings*, 2015, p. 599–608.
- [8] M. Shahid, T. Chitiboi, T. Ivanovska, Automatic mri segmentation of para-pharyngeal fat pads using interactive visual feature space analysis for classification, *BMC Med Imaging* 17 (2017). doi:10.1186/s12880-017-0179-7.
- [9] A. Ruiz, M. Rondon Sepúlveda, O. Franco, The associations between sleep disorders and anthropometric measures in adults from three colombian cities at different altitudes, *Maturitas* 94 (2016) 1–10. doi:10.1016/j.maturitas.2016.08.013.
- [10] S. Dhaliwal, S. Hesabgar, S. Haddad, Constructing a patient-specific computer model of the upper airway in sleep apnea patients, *Laryngoscope* 128 (2018) 277–282. doi:10.1002/lary.26834.
- [11] Y. Li, N. Lin, J. Ye, Upper airway fat tissue distribution in subjects with obstructive sleep apnea and its effect on retropalatal mechanical loads, *Respir Care* 57 (2012) 1098–1105. doi:10.4187/respcare.00929.
- [12] A. Kim, B. Keenan, N. Jackson, Tongue fat and its relationship to obstructive sleep apnea, *Sleep* 37 (2014) 1639–1648. doi:10.5665/sleep.4072.
- [13] S. Wang, B. Keenan, A. Wiemken, Effect of weight loss on upper airway anatomy and the apnea–hypopnea index the importance of tongue fat, *Am J Respir Crit Care Med* 201 (2020) 718–727. doi:10.1164/rccm.201903-0692OC.
- [14] F. Pérez-García, R. Sparks, S. Ourselin, Torchio: A python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning, *Comput Methods Programs Biomed* 208 (106236) (2021). doi:10.1016/j.cmpb.2021.106236.
- [15] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, O. Ronneberger, 3d u-net: Learning dense volumetric segmentation from sparse annotation,

- in: S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, W. Wells (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016, Springer International Publishing, Cham, 2016, pp. 424–432.
- [16] F. Pérez-García, fepegar/unet: Pytorch implementation of 2d and 3d u-net (2020). doi:10.5281/ZENODO.3697931.
URL <https://doi.org/10.5281/ZENODO.3697931>
- [17] A. Taha, A. Hanbury, Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool, BMC Med Imaging 15 (29) (2015). doi:10.1186/s12880-015-0068-x.
- [18] K. Zou, S. Warfield, A. Bharatha, Statistical validation of image segmentation quality based on a spatial overlap index, Acad Radiol 11 (2004) 178–189. doi:10.1016/s1076-6332(03)00671-8.
- [19] V. Bommineni, G. Erus, J. Doshi, Automatic segmentation and quantification of upper airway anatomic risk factors for obstructive sleep apnea on unprocessed magnetic resonance images. acad radiol (2022). doi:10.1016/j.acra.2022.04.023.
URL <https://doi.org/10.1016/j.acra.2022.04.023>