

Segmentación Tridimensional de Nódulos Pulmonares en Tomografía Computarizada mediante 3D U-Net utilizando al Dataset LUNA16

M. R. Ruiz Pinto
Escuela de Ingeniería
Universidad Modelo
Mérida, México

ORCID: 0009-0009-6790-8447

J. A. Sarao Sánchez
Escuela de Ingeniería
Universidad Modelo
Mérida, México
15234128@modelo.edu.mx

Abstract—Lung cancer is currently one of the main causes for cancer mortality worldwide. Lung nodules early detection by computerized tomography represents a relevant strategy for supporting early diagnosis and follow-up treatment for patients at risk. Nevertheless, identification and manual delimitation of lung nodules is a complex task due to their size, morphological variability, low contrast and its similarity with anatomical structures.

This work presents an initial phase oriented to the development of a tridimensional segmentation algorithm for lung nodules, using the public dataset LUNA16. The purpose of this phase was studying the tomography's data structure, implementing a preliminary preprocessing pipeline and evaluate the 3d U-Net architecture's behavior for delimitating nodular structures regions. The model results were evaluated by segmentation metrics such as Dice, IoU, sensibility and specificity. This phase establishes an experimental base for, in the future, complementing the system with tomographic studies from a hospital environment.

Keywords: Lung cancer, Lung nodule, Computed tomography, LUNA16, 3D U-Net, Deep Learning, Preprocessing

Resumen—El cáncer de pulmón en la actualidad sigue siendo una de las principales causas de mortalidad oncológica a nivel mundial. La detección temprana de nódulos pulmonares mediante tomografía computarizada representa una estrategia relevante para apoyar el diagnóstico temprano y el seguimiento de pacientes en riesgo. Sin embargo, la identificación y delimitación manual de nódulos pulmonares es una tarea compleja debido a su tamaño, variabilidad morfológica, bajo contraste y similitud con estructuras anatómicas.

Este trabajo presenta una fase inicial orientada al desarrollo de un algoritmo de segmentación tridimensional de nódulos pulmonares utilizando el dataset público LUNA16. El propósito de esta fase fue estudiar la estructura de los datos tomográficos, implementar un flujo inicial de preprocesamiento y evaluar el comportamiento de una arquitectura 3D U-Net para la delimitación de regiones nodulares. El desempeño del modelo fue valorado mediante métricas de segmentación como Dice, IoU, sensibilidad y especificidad. Esta etapa establece una base experimental para, posteriormente, complementar el sistema con estudios tomográficos provenientes de un entorno hospitalario.

Palabras clave: Cáncer de pulmón, Nódulos pulmonares, Tomografía computarizada, LUNA16, 3D U-Net, Aprendizaje profundo, Preprocesamiento

I. INTRODUCTION (HEADING I)

El cáncer de pulmón representa una de las enfermedades oncológicas con mayor impacto clínico debido a su alta mortalidad y a que también, muchos casos son diagnosticados en etapas avanzadas [1]. En este contexto, la tomografía computarizada de baja dosis se ha consolidado como una herramienta relevante para la selección de poblaciones con alto riesgo [2], ya que permite identificar lesiones pulmonares en etapas tempranas y realizar seguimiento de hallazgos sospechosos [3].

Dentro del análisis de estudios de tomografía computarizada torácica, uno de los intereses clínicos son los nódulos pulmonares. Un nódulo puede deberse a una alteración benigna, o potencialmente maligna; por ello, su identificación, delimitación y cuantificación volumétrica son procesos importantes dentro del radiodiagnóstico. Sin embargo, la segmentación manual de estas lesiones es una tarea demandante y dependiente del observador, especialmente cuando los nódulos son pequeños, presentan bordes poco definidos o se encuentran próximos a vasos sanguíneos, bronquios o pleura [5].

Uno de los retos técnicos principales en el procesamiento automatizado de imágenes torácicas es la similitud morfológica entre los nódulos pulmonares y ciertas estructuras anatómicas normales. En cortes axiales de CT, algunos vasos sanguíneos o bifurcaciones broncovasculares pueden presentar una apariencia similar (redonda) a la de un nódulo, lo que puede generar falsos positivos [7]. Por eso es que los sistemas de diagnóstico asistido por computadora, conocidos como CAD, buscan apoyar al especialista mediante herramientas capaces de detectar, segmentar y analizar estas regiones o zonas sospechosas dentro de grandes volúmenes de imágenes médicas [8].

Por otro lado, en estos últimos años han salido diversas técnicas de aprendizaje profundo que han mostrado avances

importantes en tareas de segmentación médica, particularmente en estudios volumétricos como la tomografía computarizada [9]. Las redes convolucionales tridimensionales permiten aprovechar la información espacial entre cortes consecutivos, lo cual es relevante para el análisis de la forma, localización y continuidad anatómica de los nódulos pulmonares [10].

La elección del dataset LUNA16 responde a la necesidad de tabajar con una base pública, estandarizada y ampliamente utilizada en investigación sobre nódulos pulmonares [12]. Gracias a estos datos se pudo estudiar la estructura de las características tomográficas y desarrollar un pipeline de procesamiento de imágenes médicas. En esta etapa, el objetivo principal del sistema no es diagnóstica si hay cáncer o clasificar malignidad, sino generar una máscara tridimensional de la región nodular y examinar su correspondencia con una referencia aproximada construida a partir de las anotaciones disponibles. Por lo cual, esta fase plantea una base metodológica inicial que posteriormente podrá complementarse con estudios tomográficos provenientes de un entorno hospitalario, con el fin de analizar el algoritmo en condiciones clínicas.

II. OBJETIVOS

1) *Objetivo General.*

Desarrollar la primera fase de un algoritmo de segmentación tridimensional de nódulos pulmonares en imágenes de tomografía computarizada utilizando el dataset público LUNA16, con el fin de establecer una base metodológica y computacional que pueda ser complementada con estudios provenientes de un entorno hospitalario.

2) *Objetivos Específicos.*

a) Caracterizar los datos tomográficos del dataset LUNA16, considerando su estructura volumétrica, formato de imagen y anotaciones disponibles.

b) Diseñar un flujo de preprocesamiento de imágenes CT que permita estandarizar los volúmenes y preparar los datos para el entrenamiento del modelo.

c) Desarrollar un modelo de segmentación tridimensional basado en una arquitectura 3D U-Net para la delimitación de zonas nodulares en tomografía computarizada.

d) Evaluar el desempeño del algoritmo de segmentación mediante métricas cuantitativas como Dice, IoU, sensibilidad y especificidad.

e) Establecer una base experimental reproducible que permita adaptar el sistema en una etapa posterior con estudios tomográficos provenientes de un entorno hospitalario.

III. ESTADO DEL ARTE

Investigaciones realizadas anteriormente han demostrado que el diagnóstico asistido por computadora es una herramienta importante para disminuir la carga de trabajo del profesional de la salud y para mitigar la variabilidad en la detección de nódulos pulmonares [1].

1) *Preprocesamiento y segmentación.*

El resampleo isotrópico a 1mm cúbico es de gran importancia para corregir la anisotropía presente en los estudios de tomografía y para asegurar la consistencia morfológica [2].

Segmentar el parénquima pulmonar funge como un filtro que delimita la región de interés y normaliza el rango de densidades en Unidades de Hounsfield entre [-1000, 400] [3], [4].

2) *Arquitectura 3D U-Net y bloques residuales.*

Se ha demostrado que las redes neuronales convolucionales 3D superan a los modelos clásicos en 2D, puesto que captura la continuidad volumétrica de las lesiones [5].

3) *Desbalance de datos y métricas.*

El estado del arte propone el entrenamiento basado en parches y el submuestreo de candidatos negativos para estabilizar el aprendizaje [1], [6].

IV. METODOLOGÍA

A. *Dataset.*

Para esta investigación, se utilizó el dataset LUNA16, que está integrado por estudios de tomografía computarizada torácica y contiene anotaciones sobre la información relacionada con los nódulos pulmonares identificados en los estudios.

Estas imágenes se encuentran en formato MHD/RAW; las anotaciones incluyen las coordenadas espaciales del centro de los nódulos detectados y el diámetro aproximado del mismo.

En este proyecto, el dataset se utilizó para diseñar un pipeline de segmentación tridimensional de nódulos pulmonares. El dataset LUNA16, no proporciona las máscaras por cada nódulo, razón por la cual se decidió generar máscaras esféricas aproximadas.

B. *Preprocesamiento de imágenes.*

Para el preprocesamiento de las imágenes se dividió el proceso en distintas etapas.

1) *Lectura de imágenes MHD/RAW.*

Para cargar las imágenes en el entorno de desarrollo, se utilizó la librería SimpleITK. El uso de esta herramienta permite conservar las características de la imagen, como lo son la información espacial del estudio, como el origen, dirección y espaciado. Posteriormente, las imágenes se convirtieron en arreglos de NumPy ordenados espacialmente como 'z,y,x'.

2) *Conversión a formato Nifti.*

Después de haber realizado el procesamiento inicial, los volúmenes se guardaron en formato Nifti; De esta manera, los datos procesados fueron organizados.

3) Resampleo isotrópico.

Los volúmenes de CT fueron remuestreados con una resolución isotrópica de 1x1x1 mm, con el objetivo de normalizar las dimensiones de todos los estudios, puesto que los diferentes estudios tomográficos pueden presentar tamaños distintos de vóxel. Uniformizar la resolución de los estudios, produce que las operaciones convoluciones 3D tengan una escala espacial consistente.

Para las imágenes tomográficas, se empleó interpolación continua, mientras que para las máscaras se utilizó la interpolación por vecino más cercano.

4) Normalización en unidades Hounsfield.

Para este paso, se definió un intervalo de -1000,400 HU. Con este rango, se puede conservar la información del parénquima pulmonar, aire, tejido blando y posibles regiones nodulares. Posteriormente, se normalizaron los valores del intervalo 0,1, con el objetivo de estabilizar el entrenamiento de la red neuronal.

5) Segmentación del parénquima pulmonar.

De igual forma, el pipeline considera incorporar máscaras del parénquima pulmonar, con el fin de restringir el análisis de la región de interés. Esta etapa, contribuye a la reducción de la influencia de estructuras externas al pulmón, para disminuir los falsos positivos.

6) Generación de máscaras de nódulos.

Utilizando las coordenadas del centro del nódulo y su diámetro aproximado, se generó una máscara esférica. Por cada uno de los nódulos, los vóxeles ubicados dentro del radio estimado fueron etiquetados como nódulo, mientras que lo demás se consideró como fondo.

7) Extracción de parches 3D.

Para el siguiente paso del desarrollo, se extrajo parches cúbicos de 64x64x64 vóxeles. Los parches positivos fueron ubicados en las regiones nodulares centrales.

8) División de datos.

Para dividir el dataset, se crearon tres subconjuntos, entrenamiento, validación y prueba, la división se realizó a nivel del paciente, con porcentajes de 70%,15% y 15% respectivamente.

C. Arquitectura 3D U-Net.

El modelo propuesto está constituido por una arquitectura 3D U-Net, implementada con MONAI/PyTorch. Esta red recibe como entrada un parche volumétrico monocanal de CT y genera en la salida una máscara tridimensional binaria.

Como configuración para la red neuronal se utilizaron los siguientes parámetros.

Hiperparámetro	Valor
Spatial Dims	3
Tamaño del parche	64*64*64
Canales internos	8,16,32,64,128
Strides	2,2,2,2
Unidades residuales	2
Normalización	Batch Normalization
Dropout	0.1

Fig. 1. Hiperparámetros utilizados en la red neuronal convolucional propuesta.

La ruta utilizada para la codificación reduce la dimensión espacial y aumenta la cantidad de mapas de características, mientras que la ruta de decodificación reconstruye la resolución del parche de salida.

D. Entrenamiento.

Para el entrenamiento del modelo, se decidió utilizar una función de pérdida compuesta por Dice Loss y Binary Cross-Entropy.

$$L=0.6L - \text{Dice} \\ +0.4L - \text{BCE}$$

Esta combinación fue escogida debido al desbalance existente entre fondo y nódulo. La mayoría de los vóxeles en un parche tridimensional corresponden a fondo, parénquima o a estructuras no nodulares, mientras que la región en la que se encuentra el nódulo ocupa una proporción considerablemente más pequeña.

Utilizar Dice Loss, favorece la superposición entre la predicción y la máscara de referencia, por otro lado, Binary Cross-Entropy contribuye al aprendizaje vóxel a vóxel.

El entrenamiento utilizó el optimizador AdamW, con una tasa de aprendizaje inicial de 1e-4, scheduler Cosine Annealing y early stopping de 20 épocas.

Durante el entrenamiento se aplicó aumento de datos tridimensionales, incluyendo rotación, inversión de ejes, ruido gaussiano y variación de intensidad. Estas transformaciones tratan de mejorar la generalización del modelo ante diferencias anatómicas.

Hiperparámetro	Valor
Épocas	30
Tasa de aprendizaje	1e-4
Optimizador	AdamW
Scheduler	CosineAnnealingLR

Hiperparámetro	Valor
Batch Size	4
Early stopping	20 épocas
Función de pérdida	combinación de 60% Dice y 40% Binary Cross-Entropy

Fig. 2. Hiperparámetros utilizados en el entrenamiento.

E. Métricas de evaluación

El modelo de segmentación fue evaluado mediante cuatro métricas.

- Coeficiente Dice

$$Dice = \frac{2TP}{2TP + FP + FN}$$

Este coeficiente mide la superposición espacial entre la máscara propuesta y la máscara de referencia. Evalúa qué tanto coincide la región segmentada por la red con la región nodular esperada.

- Intersección sobre Unión

$$IoU = \frac{TP}{TP + FP + FN}$$

La IoU evalúa la relación entre la intersección y la unión de ambas máscaras. Es un poco más estricta que Dice y permite valorar la precisión espacial de la segmentación.

- Sensibilidad

$$Sensibilidad = \frac{TP}{TP + FN}$$

La sensibilidad indica la proporción de vóxeles nodulares correctamente detectados. Esta métrica es importante ya que penaliza falsos negativos.

- Especificidad

$$\frac{TN}{TN + FP}$$

Este último coeficiente mide la capacidad del modelo para reconocer correctamente el fondo y evitar marcar regiones no nodulares como nódulos. Con una baja especificidad puede ocurrir el incremento de falsos positivos.

V. RESULTADOS.

A. Resultados cuantitativos

El modelo 3D U-Net se evaluó con el conjunto de pruebas que se utilizó mediante métricas de segmentación volumétrica. Los resultados que se obtuvieron se presentan en la Tabla I.

Tabla I. Desempeño de la 3D U-Net en el conjunto de prueba

Métrica	Resultado obtenido
Dice promedio	0.02380+/- 0.0393
IoU promedio	0.0125

Métrica	Resultado obtenido
Sensibilidad promedio	0.8393
Especificidad promedio	0.7899

Fig. 3. Resultados del test.

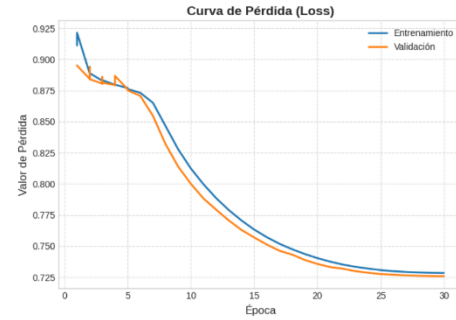


Fig. 4. Gráfica de la curva de pérdida.

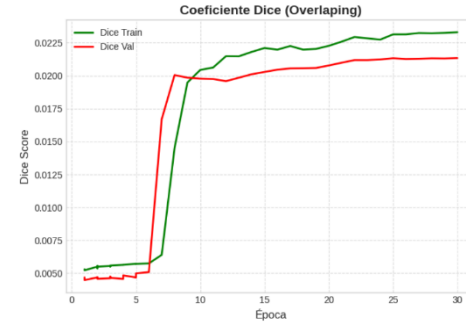


Fig. 5. Gráfica del coeficiente de Dice.

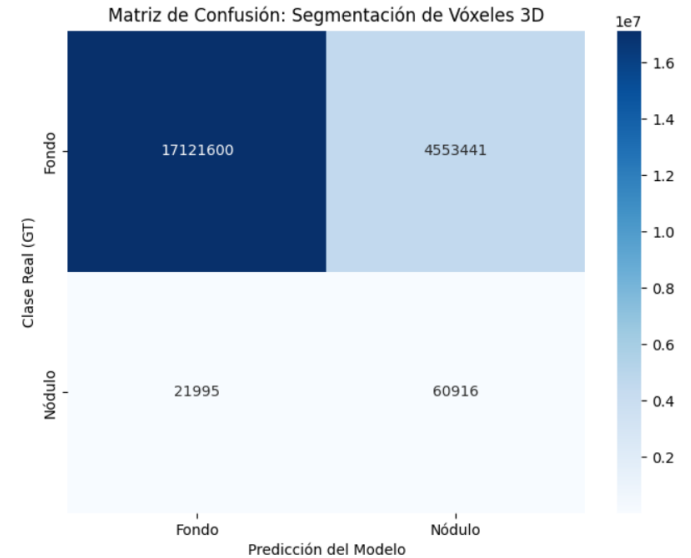


Fig. 6. Matriz de confusión.

B. Resultados cualitativos.

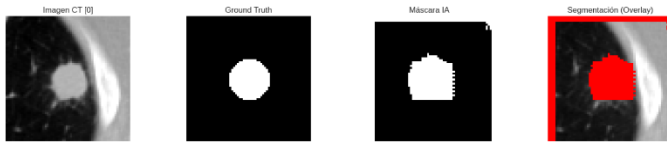


Fig. 7. Contraste entre la imagen original CT/ground truth/Máscara de la red/superposición.

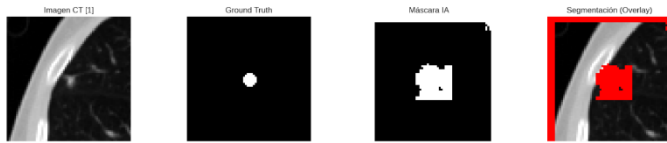


Fig. 8. Contraste entre la imagen original CT/ground truth/Máscara de la red/superposición.

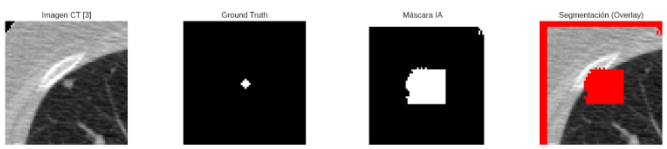


Fig. 9. Contraste entre la imagen original CT/ground truth/Máscara de la red/superposición.

VI. CONCLUSIÓN.

El trabajo permitió valorar la arquitectura 3D U-Net como una herramienta eficaz para la detección y localización de candidatos a nódulos, pero es insuficiente en su estado actual, para una delimitación diagnóstica precisa. Se logró utilizar con éxito un pipeline de preprocesamiento que nos garantiza la isotropía de los datos, lo cual es un paso importante para cualquier desarrollo posterior. Lo faltante y trabajo futuro, es que para alcanzar niveles de precisión clínica, quedan pendientes tres pilares críticos que son el refinamiento de la arquitectura, implementar mecanismos de atención para que la red ignore regiones irrelevantes del parénquima y se concentre en la morfología nodular. Estrategias de Hard Negative Mining: Es necesario forzar a la red a entrenar con los casos donde más falla (vasos sanguíneos y cicatrices), reduciendo así los falsos positivos. En conclusión, los resultados establecen una "línea base" funcional que demuestra el potencial del aprendizaje profundo 3D, subrayando la necesidad de optimizar las funciones de costo y aumentar el tiempo de entrenamiento para mejorar la especificidad del sistema.

VII. DISCUSIÓN.

Los resultados que fueron obtenidos reflejan los desafíos que representa la segmentación volumétrica de lesiones de tamaño reducido. El coeficiente de Dice resultante de 0.0238 remarca una precisión geométrica limitada, atribuida principalmente a

que el modelo logra localizar la presencia del nódulo, pero presenta complicaciones para delimitar con exactitud los bordes de la lesión. En la mayoría de los casos, el modelo subsegmenta o genera falsos positivos por las estructuras vasculares adyacentes debido a la similitud en unidades de Hounsfield, no obstante, es un precedente para el procesamiento de las imágenes de tomografía y funge como base para las etapas posteriores del proyecto.

REFERENCIAS

- [1] F. Bray, M. Laversanne, H. Sung, J. Ferlay, R. L. Siegel, I. Soerjomataram, and A. Jemal, "Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 74, no. 3, pp. 229–263, 2024.
- [2] A. H. Krist, K. W. Davidson, C. M. Mangione, M. J. Barry, M. Cabana, A. B. Caughey, et al., "Screening for lung cancer: US Preventive Services Task Force recommendation statement," *JAMA*, vol. 325, no. 10, pp. 962–970, 2021.
- [3] P. J. Mazzone, G. A. Silvestri, L. H. Souter, T. J. Caverly, J. P. Kanne, H. A. Katki, et al., "Screening for lung cancer: CHEST guideline and expert panel report," *Chest*, vol. 160, no. 5, pp. e427–e494, 2021.
- [4] S. F. Banu, M. M. K. Sarker, M. Abdel-Nasser, D. Puig, and H. A. Rashwan, "AWEU-Net: An attention-aware weight excitation U-Net for lung nodule segmentation," *Applied Sciences*, vol. 11, no. 21, Art. no. 10132, 2021.
- [5] D. Lu, J. Chu, R. Zhao, Y. Zhang, and G. Tian, "A novel deep learning network and its application for pulmonary nodule segmentation," *Computational Intelligence and Neuroscience*, vol. 2022, Art. no. 7124902, 2022.
- [6] I. Marinakis, K. Karampidis, and G. Papadourakis, "Pulmonary nodule detection, segmentation and classification using deep learning: A comprehensive literature review," *BioMedInformatics*, vol. 4, no. 3, pp. 2043–2106, 2024.
- [7] L. Zhi, W. Jiang, S. Zhang, and T. Zhou, "Deep neural network pulmonary nodule segmentation methods for CT images: Literature review and experimental comparisons," *Computers in Biology and Medicine*, vol. 164, Art. no. 107321, 2023.
- [8] Z. Ji, Z. Zhao, X. Zeng, J. Wang, L. Zhao, X. Zhang, and I. Ganchev, "ResDSda U-Net: A novel U-Net-based residual network for segmentation of pulmonary nodules in lung CT images," *IEEE Access*, vol. 11, pp. 87775–87789, 2023.
- [9] M. J. Cardoso, W. Li, R. Brown, N. Ma, E. Kerfoot, Y. Wang, et al., "MONAI: An open-source framework for deep learning in healthcare," *arXiv preprint arXiv:2211.02701*, 2022.
- [10] T. Hu, Y. Lan, Y. Zhang, J. Xu, S. Li, and C.-C. Hung, "A lung nodule segmentation model based on the transformer with multiple thresholds and coordinate attention," *Scientific Reports*, vol. 14, Art. no. 31743, 2024.
- [10] T. Tang and R. Zhang, "A multi-task model for pulmonary nodule segmentation and classification," *Journal of Imaging*, vol. 10, no. 9, Art. no. 234, 2024.
- [11] [X. Ma, H. Song, X. Jia, and Z. Wang, "An improved V-Net lung nodule segmentation model based on pixel threshold separation and attention mechanism," *Scientific Reports*, vol. 14, Art. no. 4743, 2024.
- [13] T. M. M. Aung and A. A. Khan, "Enhanced U-Net with attention mechanisms for improved feature representation in lung nodule segmentation," *Current Medical Imaging*, vol. 21, 2025.