

CLASIFICACIÓN DE LOS NÓDULOS PULMONARES UTILIZANDO  
REPRESENTACIONES PROFUNDAS DEDICADAS A LA SEGMENTACIÓN Y  
CARACTERIZACIÓN DE FORMA.

LAURA VANESSA MÉNDEZ RIVERA

UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FISICOMECAÑICAS  
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA  
BUCARAMANGA

2025

CLASIFICACIÓN DE LOS NÓDULOS PULMONARES UTILIZANDO  
REPRESENTACIONES PROFUNDAS DEDICADAS A LA SEGMENTACIÓN Y  
CARACTERIZACIÓN DE FORMA.

Laura Vanessa Méndez Rivera

Trabajo de Grado presentado en cumplimiento de los requisitos para optar al título de:  
Ingeniera de Sistemas

Director:

Fabio Martínez Carrillo

Doctor en Ingeniería de Sistemas y Computación

Codirectora:

Alejandra Moreno Tarazona

Magíster en Ingeniería de Sistemas e Informática

UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FISICOMECAÑICAS  
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA  
BUCARAMANGA

2025

## **AGRADECIMIENTOS**

Primero, agradezco a Dios por guiar cada paso de mi camino, por darme la fuerza en los momentos difíciles y por permitirme culminar esta etapa tan importante de mi vida.

A mi familia, gracias por ser el pilar fundamental en mi formación. A mis padres, por su amor incondicional, por estar siempre presentes, por apoyarme en cada decisión y por enseñarme con el ejemplo. A mis hermanos, Karina y Anderson, gracias por ser modelos de dedicación y constancia, por inspirarme y acompañarme. A mis cuñados, Dayana y Fabián, y especialmente a mi sobrina Anita, quien con su llegada llenó de alegría los últimos momentos de esta etapa; gracias por traer luz y ternura cuando más lo necesitaba. A toda mi familia: a mis nonitos, mis tías, primos y a mi tío Alfonso, gracias por su cariño y por estar siempre pendientes de mí y de mi universidad. A mi fiel compañía, Muñeca, gracias por trasnochar conmigo tantas veces, por estar a mi lado en silenciosa lealtad mientras trabajaba.

A mi director, Fabio, gracias por su tiempo, por su paciencia, su dedicación y por creer en este proyecto. A mi codirectora y amiga Alejandra, gracias por sus consejos, por su acompañamiento constante y por compartir conmigo no solo su conocimiento, sino también su amistad.

A José Daniel, mi pareja, gracias por caminar conmigo este proceso, por ser apoyo, motivación y presencia constante. A mis compañeros de BIVL<sup>2</sup>ab, gracias por su ayuda, por el compañerismo, por las tardes de risas, por cada reunión y seminario donde estuvieron presentes, aconsejándome y animándome a seguir.

Finalmente, a mis amigos de la carrera, gracias por los momentos compartidos, por las

conversaciones que calman y motivan, por crecer juntos en este camino lleno de aprendizajes y desafíos. Llevo conmigo el cariño de todos ustedes y el valor de cada experiencia vivida.

## CONTENIDO

	<b>pág.</b>
<b>INTRODUCCIÓN</b> . . . . .	<b>10</b>
<b>1. FUNDAMENTOS Y TRABAJOS PREVIOS</b> . . . . .	<b>14</b>
1.1. Cáncer de pulmón . . . . .	14
1.2. Métodos computacionales de segmentación. . . . .	17
1.3. Caracterización de nódulos pulmonares desde enfoques computacionales. . .	19
<b>2. PROBLEMA DE INVESTIGACIÓN</b> . . . . .	<b>23</b>
<b>3. OBJETIVOS</b> . . . . .	<b>24</b>
3.1. Objetivo general . . . . .	24
3.2. Objetivos específicos . . . . .	24
<b>4. MÉTODO PROPUESTO</b> . . . . .	<b>25</b>
4.1. Autoencoder para determinación de forma nodular . . . . .	26
4.2. Descriptores geométricos . . . . .	29
4.3. Clasificación de descriptor de forma . . . . .	30
4.4. Diseño Experimental . . . . .	32
<b>5. EVALUACIÓN Y RESULTADOS</b> . . . . .	<b>38</b>
<b>6. CONCLUSIONES Y TRABAJO FUTURO</b> . . . . .	<b>44</b>
<b>BIBLIOGRAFÍA</b> . . . . .	<b>48</b>

## LISTA DE FIGURAS

	pág.
Figura 1. Diagrama de la arquitectura U-Net para segmentación de nódulos pulmonares. <sup>1</sup> . . . . .	18
Figura 2. Esquema de una Red Neuronal Convolutiva Totalmente Conectada (FCN) <sup>2</sup>	19
Figura 3. Esquema general de la estrategia implementada, donde se integran descriptores aprendidos y anotaciones radiológicas. B) Detalle del módulo de atención implementado $\beta$ . C) Extracción de los descriptores geométricos calculados a partir de las segmentaciones generadas por el modelo. D) Validación mediante un clasificador Random Forest. . . . .	25
Figura 4. Ejemplos de segmentaciones predichas por la U-Net en diferentes casos de nódulos pulmonares. . . . .	38
Figura 5. Comparación de la distribución del embebido generado por la U-Net con atención (1) y por el método propuesto (2). . . . .	43

---

<sup>1</sup> Olaf RONNEBERGER; Philipp FISCHER, and Thomas BROX. “U-net: Convolutional networks for biomedical image segmentation”. In: *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer. 2015, pp. 234–241.

<sup>2</sup> Evan SHELHAMER; Jonathan LONG, and Trevor DARRELL. “Fully convolutional networks for semantic segmentation”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.4 (2017), pp. 640–651.

## LISTA DE TABLAS

	<b>pág.</b>
Tabla 1. Descripción de los tipos de nódulos pulmonares. <sup>34</sup> . . . . .	16
Tabla 2. Resultados del Dice Score y Hausdorff para las diferentes configuraciones de la U-Net. . . . .	39
Tabla 3. Evaluación individual del desempeño de los descriptores geométricos y radiológicos utilizados como entrada al clasificador, en comparación con el método propuesto que combina ambos conjuntos de características. . . . .	40
Tabla 4. Comparación del desempeño de diferentes configuraciones de U-Net: línea base, con mecanismo de atención, integrando descriptores geométricos o morfológicos radiológicos, y el método propuesto que combina ambas fuentes de información. . . . .	41
Tabla 5. Comparación del método propuesto con enfoques del estado del arte en la clasificación de nódulos pulmonares. . . . .	42

---

<sup>3</sup> Nichole T TANNER et al. “Management of pulmonary nodules by community pulmonologists: a multi-center observational study”. In: *Chest* 148.6 (2015), pp. 1405–1414.

<sup>4</sup> Ashleigh CRUICKSHANK; Geoff STIELER, and Faisal AMEER. “Evaluation of the solitary pulmonary nodule”. In: *Internal Medicine Journal* 49.3 (2019), pp. 306–315.

## RESUMEN

**TÍTULO:** Clasificación de los nódulos pulmonares utilizando representaciones profundas dedicadas a la segmentación y caracterización de forma. \*

**AUTOR:** Laura Vanessa Méndez Rivera \*\*

**PALABRAS CLAVE:** Nódulos pulmonares, tomografía computarizada, aprendizaje débilmente supervisado, métodos discriminativos, segmentación, clasificación.

**DESCRIPCIÓN:** El cáncer de pulmón presenta la mayor incidencia y mortalidad relacionada con el cáncer a nivel mundial; en 2022, aproximadamente el 18.7% de la población falleció por esta causa. El diagnóstico se basa en el análisis de nódulos pulmonares, pequeñas masas anormales consideradas precursores de la enfermedad y principales biomarcadores de las mismas, identificadas mediante tomografías computarizadas (TC) de baja dosis que permiten evaluar textura, tamaño, densidad y morfología. No obstante, su diagnóstico es subjetivo, reportándose un acuerdo bajo/moderado (kappa promedio de 0.51 y 0.57 para la variabilidad inter-intra-observador, respectivamente). En la actualidad, se han desarrollado diversas metodologías para la clasificación de los nódulos pulmonares, pero siguiendo estrategias supervisadas, con dependencia de anotaciones de expertos. Por tanto, este trabajo presenta una estrategia que permite realizar la clasificación de los nódulos pulmonares utilizando representaciones profundas dedicadas a la segmentación y caracterización de la forma. Para ello, se implementó una U-Net con mecanismos de atención que permiten extraer características densas, enfatizando bordes y texturas. Adicionalmente, se integraron tokens radiológicos en el proceso, los cuales se concatenaron con el embebido generado por la U-Net, combinando información sobre lobulación, espiculación, textura y margen. Posteriormente, este embebido combinado se enriqueció al concatenarse con las características geométricas obtenidas a partir de las segmentaciones, como el área y el perímetro, logrando una representación final más robusta para la clasificación. El método fue evaluado en conjuntos de datos públicos, alcanzando un AUC de 93.02%  $\pm$ 0.0018, una sensibilidad de 94.86%  $\pm$ 0.010, una precisión de 95.79%  $\pm$ 0.005 y un F1-score de 95.32%  $\pm$ 0.005, lo que demuestra su potencial para el análisis y clasificación de nódulos pulmonares.

---

\* Trabajo de investigación

\*\* Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería de Sistemas e Informática. Director: Fabio Martínez, PhD en ingeniería de sistemas y computación, análisis de imágenes y análisis de vídeo. Codirectora: Alejandra Moreno Tarazona, Magíster en Ingeniería de Sistemas e Informática

## ABSTRACT

**TITLE:** Lung nodule classification by using deep representations dedicated to segmentation and shape characterization \* :

**AUTHOR:** Laura Vanessa Méndez Rivera \*\*

**KEYWORDS:** Pulmonary nodules, computed tomography, weakly supervised learning, discriminative methods, segmentation, classification.

**DESCRIPTION:** Lung cancer has the highest incidence and cancer-related mortality worldwide; in 2022, approximately 18.7% of the population died from this cause. Diagnosis is based on the analysis of pulmonary nodules—small abnormal masses considered precursors of the disease and key biomarkers—identified through low-dose computed tomography (CT), which allows for the assessment of texture, size, density, and morphology. However, the diagnosis is subjective, reporting a low/moderate agreement (average kappa of 0.51 and 0.57 for inter-intra-observer variability, respectively). Currently, several computational methodologies have been developed allowing the classification of lung nodules, mainly between extreme stages (benign vs malignant). However, these strategies are supervised and annotation-dependent by experts. The proposed methodology presents a strategy to classify pulmonary nodules using deep representations dedicated to segmentation and morphological characterization. A U-Net with attention mechanisms is employed, in order to extract dense features, enhancing edges, and textures. Additionally, radiological tokens were integrated into the process and concatenated with the U-Net embedding, incorporating information on lobulation, spiculation, texture, and margins. The combined embedding was further enriched by incorporating geometric features derived from the segmentations, such as area and perimeter, resulting in a more robust representation for classification. The proposed method was evaluated on public datasets by achieving an AUC of 93.02%  $\pm$ 0.005, a recall of 94.86%  $\pm$ 0.010, a precision of 95.79%  $\pm$ 0.005, and an F1-score of 95.32%  $\pm$ 0.005, demonstrating its potential for pulmonary nodule analysis and classification.

---

\* Research work

\*\* Faculty of Physics-Mechanics Engineering. School of Systems Engineering and Informatics. Advisor: Fabio Martínez Carrillo, PhD. Computer and systems engineering, medical image analysis, and video analysis. Co-advisor: Alejandra Moreno Tarazona, Master in Systems and Informatics Engineering.

## INTRODUCCIÓN

El cáncer de pulmón es la principal causa de muerte relacionada con el cáncer a nivel mundial, reportando 1.8 millones de personas fallecidas en el 2022<sup>1</sup>. De igual manera es el cáncer con mayor incidencia a nivel global con  $\sim 12.4\%$  casos registrados<sup>1</sup>. A nivel nacional, estas cifras no son atípicas, siendo este el segundo cáncer con mayor tasa de mortalidad, representando aproximadamente el 10.8% de los casos<sup>1</sup>.

Los nódulos pulmonares son masas pequeñas o áreas anormales que varían entre los [3 – 30] milímetros y se consideran uno de los principales biomarcadores para la localización y el diagnóstico del cáncer pulmonar<sup>2</sup>. En la clínica, estos nódulos son delineados y estratificados mediante tomografías computarizadas (CT, por sus siglas en inglés) de baja dosis, los cuales se consideran la modalidad de imagen más sensible para la detección de nódulos pulmonares<sup>3</sup>. Con la finalidad de realizar una clasificación de la malignidad de estas masas, generalmente se tienen en cuenta diversas características tales como la morfología, la textura, el tamaño, y la densidad<sup>4</sup>. Sin embargo, el diagnóstico de la malignidad es subjetivo, presentando un acuerdo entre bajo y moderado ( $\kappa_{inter} = 0.51, \kappa_{intra} = 0.57$ ) tanto intra como inter-observador, demostrando la variabilidad

- 
- <sup>1</sup> INTERNATIONAL AGENCY FOR RESEARCH ON CANCER. *Global Cancer Observatory*. <https://gco.iarc.fr/>. 2022.
  - <sup>2</sup> Jorge Juan Suárez CUENCA. “Desarrollo de un sistema de diagnóstico asistido por computador para detección de nódulos pulmonares en tomografía computarizada multicorte.” In: (2009).
  - <sup>3</sup> “Estudio y Seguimiento del Nódulo Pulmonar Solitario Detectado por RX y TC”. in: (2023).
  - <sup>4</sup> Kai-Lung HUA et al. “Computer-aided classification of lung nodules on computed tomography images via deep learning technique”. In: *OncoTargets and therapy* (2015), pp. 2015–2022.

del problema <sup>5</sup>. De la misma forma, en términos de segmentación, sus tasas no son tan diferentes, dado que presenta un acuerdo moderado ( $\kappa_{Fleiss} = 0.65$ ), dificultando la caracterización de los nódulos pulmonares.

En la literatura se han desarrollado múltiples estrategias computacionales que buscan realizar la clasificación de nódulos pulmonares a través de redes convolucionales profundas<sup>6 7</sup>. Generalmente, existen metodologías basadas en convoluciones que extraen patrones de textura y bordes en las imágenes, lo cual puede facilitar la clasificación de los nódulos. También existen otras propuestas que integran mecanismos de atención en redes convolucionales, permitiendo a los modelos enfocar su análisis en áreas relevantes de la imagen y capturar relaciones espaciales de manera no local. Estos mecanismos de atención amplían el campo receptivo y mejoran la capacidad del modelo para identificar patrones globales y sutiles<sup>8 9</sup>. También se han reportado en la literatura, métodos basados en radiómica y morfología que han permitido incorporar características derivadas de la forma, tamaño y textura del nódulo, aportando información adicional. Estas características ofrecen una representación más detallada y estructurada, que refuerza el análisis

---

<sup>5</sup> Sarah J van RIEL et al. “Observer variability for classification of pulmonary nodules on low-dose CT images and its effect on nodule management”. In: *Radiology* 277.3 (2015), pp. 863–871.

<sup>6</sup> Xinzhuo ZHAO et al. “Agile convolutional neural network for pulmonary nodule classification using CT images”. In: *International journal of computer assisted radiology and surgery* 13 (2018), pp. 585–595.

<sup>7</sup> Sozan Abdullah MAHMOOD and Hunar Abubakir AHMED. “An improved CNN-based architecture for automatic lung nodule classification”. In: *Medical & Biological Engineering & Computing* 60.7 (2022), pp. 1977–1986.

<sup>8</sup> Sarfaraz HUSSEIN et al. “Tumornet: Lung nodule characterization using multi-view convolutional neural network with gaussian process”. In: *2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017)*. IEEE. 2017, pp. 1007–1010.

<sup>9</sup> Xiaohang FU et al. “Attention-enhanced cross-task network for analysing multiple attributes of lung nodules in ct”. In: *arXiv preprint arXiv:2103.03931* (2021).

de los nódulos a partir de su geometría <sup>10 11</sup>. No obstante, estos métodos continúan acotados a esquemas supervisados, orientados a etiquetas de malignidad. De hecho, los patrones morfológicos aprendidos son obtenidos desde reglas de minimización de la malignidad, lo que puede limitar el margen de variabilidad de las observaciones, así como también puede sesgarse por anotaciones de malignidad corruptas.

El presente trabajo implementó una metodología para el modelamiento morfológico de los nódulos pulmonares que pudiesen servir como descriptores débilmente supervisados en la tarea de estratificación de la malignidad. Primero, se implementó una arquitectura auto-codificadora que recibe observaciones nodulares y se dedica a obtener su segmentación. Para ello se implementó una red U-Net, reforzada con mecanismos de atención, que además recibe información estratificada relacionada con características radiológicas marcadas por radiólogos (*LUNG-RADS tokens*) para guiar la delineación automática de los nódulos, preservando así características locales, útiles en la definición de la malignidad. Estos *tokens* son integrados mapeando la información tabular en capas densas, que se concatenan al vector embebido de la representación visual. Desde esta arquitectura se obtienen vectores embebidos que actúan como descriptores de nódulos, recopilando información textural y morfológica, codificada para realizar la segmentación. También desde la segmentación obtenida se extraen el área y perímetro, la relación de aspecto, la solidez, entre otras como medidas complementarias de forma nodular. Entonces, tanto el vector embebido como las medidas de la segmentación son concatenadas para formar un descriptor nodular el cual es mapeado a un clasificador para obtener una medida de malignidad. El trabajo fue entrenado y validado en un conjunto de datos públicos

---

<sup>10</sup> Qing-qing XU et al. "Prediction efficacy of feature classification of solitary pulmonary nodules based on CT radiomics". In: *European Journal of Radiology* 139 (2021), p. 109667.

<sup>11</sup> Mehdi ASTARAKI et al. "Benign-malignant pulmonary nodule classification in low-dose CT with convolutional features". In: *Physica Medica* 83 (2021), pp. 146–153.

con 530 nódulos, que contenían información relacionada con la anotación y diagnóstico de la malignidad a través de múltiples radiólogos expertos en el tema de los nódulos pulmonares.

## 1. FUNDAMENTOS Y TRABAJOS PREVIOS

### 1.1. Cáncer de pulmón

El cáncer de pulmón representa la principal causa de mortalidad por cáncer a nivel global. En 2022, se estima que alrededor de 1.8 millones de personas fallecieron por causa de esta enfermedad<sup>1</sup>. El diagnóstico temprano es clave para reducir la mortalidad por esta enfermedad y para mitigar efectos adversos en los pacientes. Los nódulos pulmonares son masas anormales con tamaños que oscilan entre [3-30] milímetros, considerados como los precursores del cáncer de pulmón y uno de los principales biomarcadores del cáncer de pulmón. Con la finalidad de realizar una detección y diagnóstico de los nódulos pulmonares, se utilizan tomografías computacionales (CT) de baja dosis, que permiten visualizar la morfología de los nódulos pulmonares, delimitando las masas anormales y permitiendo una clasificación de la malignidad. En términos radiológicos, el diagnóstico de los nódulos pulmonares se centra en varios aspectos fundamentales, tales como el tamaño, textura, morfología, densidad, y localización<sup>12</sup>. En cuanto al tamaño los nódulos pequeños entre [3 – 8] mm tienen una probabilidad de 0.3% de ser malignos<sup>13</sup>. En contraste, la tasa de malignidad de los nódulos mayores a 10 mm incrementa considerablemente, siendo de 15.2%<sup>14</sup>. Adicionalmente, aquellos nódulos grandes, los cuales miden entre [21 – 30] mm exhiben una probabilidad aún mayor, alcanzando el 17.5%<sup>13</sup>. También los expertos radiólogos realizan una caracterización de la textura y morfología,

---

<sup>12</sup> Cecilia Irene Loeza MEJIA; Rajesh Roshan BISWAL, and Eddy SÁNCHEZ-DELACRUZ. “Estado del arte de clasificación de nódulos pulmonares utilizando Aprendizaje Automático”. In: (2019).

<sup>13</sup> Denise R ABERLE et al. “Results of the two incidence screenings in the National Lung Screening Trial”. In: *New England Journal of Medicine* 369.10 (2013), pp. 920–931.

<sup>14</sup> Ashleigh CRUICKSHANK; Geoff STIELER, and Faisal AMEER. “Evaluation of the solitary pulmonary nodule”. In: *Internal Medicine Journal* 49.3 (2019), pp. 306–315.

evaluando criterios como la calcificación, presencia de tejido adiposo y esfericidad. También, la espiculación o lobulación, están asociadas con un considerable aumento en el riesgo de malignidad, presentando una tasa de malignidad del 94.07%<sup>15</sup>. En la Tabla 1 se detallan las características específicas de los nódulos pulmonares.

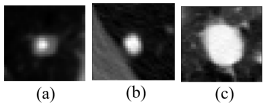







Con la finalidad de estandarizar la evaluación del diagnóstico pulmonar, se han desarrollado protocolos como Lung-RADS y el de la Sociedad Fleischner los cuales buscan minimizar la confusión en la interpretación de CT y mejorar el seguimiento de los pacientes, optimizando así la calidad del diagnóstico y tratamiento de los nódulos pulmonares. Específicamente, el protocolo Lung-RADS permite diagnosticar y realizar el seguimiento de los nódulos pulmonares detectados en las pruebas de cribado del cáncer de pulmón mediante CT de baja dosis. Este protocolo permite definir un conjunto de características geométricas y morfológicas que logran determinar y caracterizar los estadios de los nódulos pulmonares. Además de ello, asigna una categoría la cual está entre [0 – 4], esta indica la probabilidad de malignidad, la prevalencia que puede poseer y la recomendación de seguimiento que debe tener el paciente. A medida que el número incrementa, mayor es el riesgo de cáncer pulmonar<sup>18</sup>.

De manera similar, la Sociedad Fleischner, una sociedad internacional multidisciplinaria que se dedica al diagnóstico y tratamiento de las enfermedades del tórax, ha establecido unas pautas para brindar una estandarización en la evaluación de los nódulos pul-

---

<sup>15</sup> Li YANG; Q. ZHANG, and et. AL. “Assessment of the cancer risk factors of solitary pulmonary nodules”. In: *Oncotarget* 8.17 (2017), p. 29318.

<sup>18</sup> Julia KASTNER et al. “Lung-RADS version 1.0 versus lung-RADS version 1.1: comparison of categories using nodules from the national lung screening trial”. In: *Radiology* 300.1 (2021), pp. 199–206.

Características	Representación	Ejemplos
Tamaño	Dimensiones físicas del nódulo, usualmente medidas en milímetros. Los tipos de tamaños son: pequeño (3-8mm)[a], mediano (8-12mm)[b], y grande (12-30mm)[c].	
Espiculación	Presencia de bordes irregulares o espiculados. Su probabilidad de ser malignos incrementa entre el 20% y el 90%.	
Lobulación	Contornos que presentan protuberancias en la estructura del nódulo. La probabilidad de malignidad varía entre el 20% y el 95%.	
Esfericidad	Grado en que un nódulo se asemeja a una forma esférica o redondeada. La probabilidad de malignidad varía entre el 5% y el 10%.	
Calcificaciones	Depósitos de minerales, principalmente calcio, que se forman dentro del tejido del nódulo. La probabilidad de malignidad oscila entre el 8% y el 35%.	
Nódulos sólidos	Nódulos con una densidad similar o mayor que el parénquima pulmonar. Tienen una mayor probabilidad de ser malignos en comparación con los no sólidos oscilando entre el 10% y el 60%.	
Nódulos no sólidos	Nódulos con una densidad menor que el parénquima pulmonar. Generalmente la probabilidad de malignidad varía entre el 5% y el 25%.	
Nódulos mixtos	Nódulos que se caracterizan por tener áreas de densidad sólida combinada con áreas de atenuación en vidrio deslustrado. La malignidad incrementa, especialmente si el componente sólido crece, en donde su probabilidad se encuentra entre el 20% y el 60%.	

**Tabla 1.** Descripción de los tipos de nódulos pulmonares.<sup>1617</sup>

monares detectados de forma incidental<sup>19</sup>. Las pautas de la Sociedad Fleischner ofrecen

<sup>19</sup> Tamara Laxe VIDAL et al. "Enfrentándonos a la variabilidad: recomendaciones de la Sociedad Fleischner para medir correctamente los nódulos pulmonares". In: *Seram* 1.1 (2021).

recomendaciones basadas en la evidencia y la experiencia clínica para el seguimiento, la evaluación y el tratamiento de los nódulos pulmonares, clasificándolos en sólidos y subsólidos, teniendo en cuenta el tamaño, la morfología, la ubicación y el riesgo de malignidad <sup>19</sup>.

## 1.2. Métodos computacionales de segmentación.

Hoy en día, los métodos de segmentación se basan en representaciones profundas, entrenadas con etiquetas de delineación de la forma de interés, brindando resultados prometedores en diversas áreas y aplicaciones, incluido la segmentación de los nódulos pulmonares <sup>20</sup>. Entre estas técnicas, las arquitecturas de redes neuronales como U-Net y las FCN se destacan por su capacidad para capturar la morfología de los nódulos, identificando características como bordes, texturas y formas. Esto es crucial ya que atributos, como la forma de un nódulo, pueden ser indicativos de su naturaleza benigna o maligna, lo cual es esencial para un diagnóstico correcto del cáncer de pulmón <sup>21</sup>.

La arquitectura U-Net <sup>22</sup> figura 1, opera mediante un enfoque de codificador-decodificador. El codificador reduce progresivamente la resolución espacial de la imagen mientras extrae características importantes. Por otro lado, el decodificador incrementa la resolución de las características codificadas utilizando convoluciones transpuestas, crucial para reconstruir la imagen a su resolución original. Un componente esencial de la U-Net son las

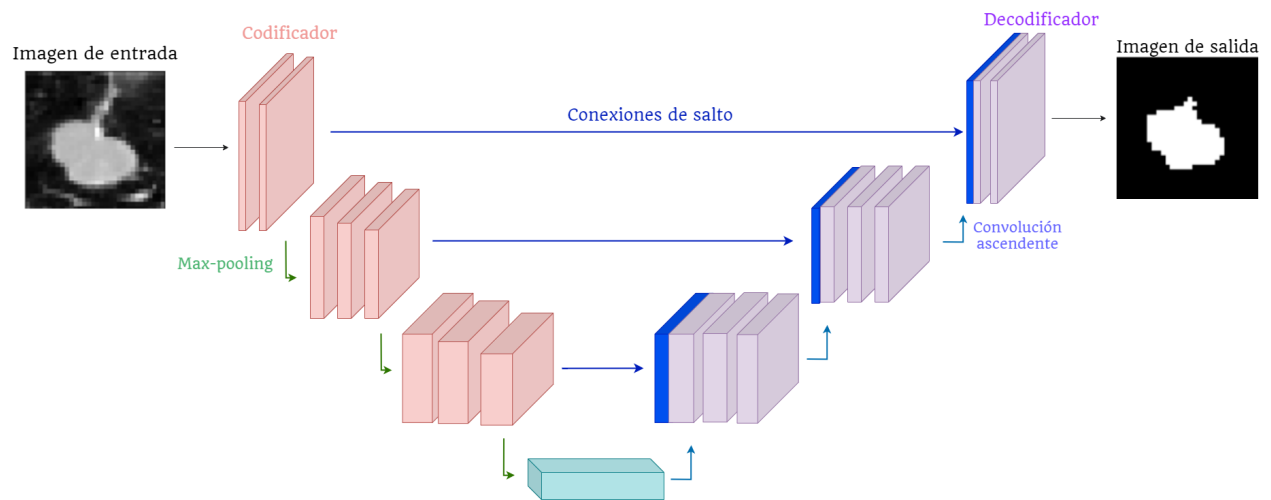
---

<sup>20</sup> Hui YU et al. "Design of lung nodules segmentation and recognition algorithm based on deep learning". In: *BMC bioinformatics* 22 (2021), pp. 1–21.

<sup>21</sup> Guofeng TONG et al. "Improved U-NET network for pulmonary nodules segmentation". In: *Optik* 174 (2018), pp. 460–469.

<sup>22</sup> Olaf RONNEBERGER; Philipp FISCHER, and Thomas BROX. "U-net: Convolutional networks for biomedical image segmentation". In: *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer. 2015, pp. 234–241.

conexiones de salto *skip connections*. Estas conexiones permiten transferir directamente la información espacial desde las capas de bajada *downsampling* del codificador a las capas correspondientes de *upsampling* en el decodificador. Al transferir esta información espacial de alta resolución, el modelo puede recuperar detalles finos y preservar los bordes en la imagen segmentada, mejorando la precisión del modelo en tareas de segmentación médica.

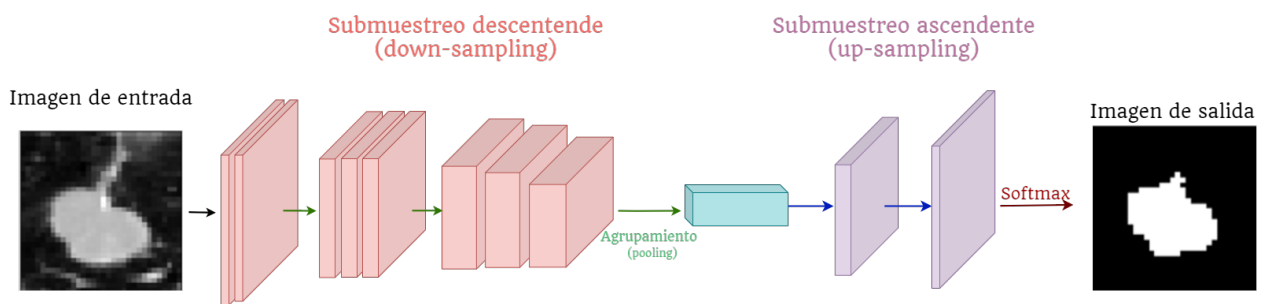


**Figura 1.** Diagrama de la arquitectura U-Net para segmentación de nódulos pulmonares.<sup>23</sup>

Así mismo, las FCN también se han diseñado para abordar las necesidades de la segmentación médica, como en la detección de nódulos pulmonares, gracias a su capacidad para manejar entradas de dimensiones variables tal y como se puede visualizar en la Figura 2. Esta flexibilidad se logra eliminando las capas completamente conectadas, permitiendo un procesamiento más eficiente y adaptativo de las variadas dimensiones de las imágenes médicas <sup>24</sup>. En las FCN, las imágenes pasan a través de capas convolucionales que reducen progresivamente su resolución mientras extraen patrones, tales como bordes y texturas de los nódulos pulmonares. Este proceso de extracción de car-

<sup>24</sup> Evan SHELHAMER; Jonathan LONG, and Trevor DARRELL. “Fully convolutional networks for semantic segmentation”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.4 (2017), pp. 640–651.

acterísticas es crucial para identificar detalles relevantes en la imagen. A medida que la imagen se procesa, las dimensiones espaciales se reducen, pero la información esencial sobre los patrones presentes se preserva y se condensa en un espacio más pequeño. Posteriormente, las FCN implementan técnicas de *up-sampling*, lo que permite aumentar la resolución de la imagen procesada. Este *up-sampling* es esencial para mantener la precisión en la segmentación a nivel del píxel, asegurando que los detalles finos y las estructuras no se pierdan durante el proceso. El proceso culmina con la clasificación de cada píxel en la imagen. Para lograr esto, se emplea una función *Softmax* aplicada sobre las características combinadas. Esta función asigna una probabilidad a cada clase posible para cada píxel, permitiendo determinar con precisión a qué clase pertenece cada píxel en la imagen segmentada.



**Figura 2.** Esquema de una Red Neuronal Convolutiva Totalmente Conectada (FCN) <sup>25</sup>

### 1.3. Caracterización de nódulos pulmonares desde enfoques computacionales.

En los últimos años, se han propuesto diferentes estrategias computacionales que permiten caracterizar, clasificar y segmentar los nódulos pulmonares detectados en estudios de CT por medio de la forma<sup>26</sup> <sup>27</sup>. En cuanto a la clasificación por ejemplo, Penghua

<sup>26</sup> Diego RIQUELME and Moulay A AKHLOUFI. “Deep learning for lung cancer nodules detection and classification in CT scans”. In: *Ai* 1.1 (2020), pp. 28–67.

<sup>27</sup> Rui LI et al. “Deep learning applications in computed tomography images for pulmonary nodule detection and diagnosis: A review”. In: *Diagnostics* 12.2 (2022), p. 298.

Zhai *et al.* han implementado una red neuronal convolucional multitarea (MT-CNN), que procesa múltiples vistas en 2D de cada nódulo a través de modelos individuales. Esta red utiliza una combinación de ramas de clasificación y reconstrucción de imágenes para disminuir la tasa de falsos positivos y aumentar la precisión en la identificación de nódulos malignos <sup>28</sup>. Por otro lado, Asghar Ali Shah *et al.* han desarrollado una red convolucional ensamblada que permite detectar nódulos pulmonares dependiendo de características tales como la forma y textura de los mismo. Este modelo usa tres CNNs que permiten caracterizar y cuantificar los nódulos pulmonares permitiendo clasificarlos entre nódulo y no nódulo <sup>29</sup>. Sin embargo, estos enfoques no integran descriptores de forma detallada en el embebido final para la clasificación de la malignidad y son estrictamente supervisados, lo que puede conducir a sesgos en las representaciones.

Metodologías en 3D han permitido capturar detalles estructurales o de bordes de los nódulos pulmonares a través de múltiples direcciones, como por ejemplo, Chao Tong *et al.* que propuso clasificar nódulos desde redes neuronales convolucionales tridimensionales (3D-CNN) y máquinas de vector de soporte (SVM) con algoritmos de aprendizaje de múltiples núcleos (MKL). Este sistema captura detalles estructurales del nódulo desde múltiples direcciones en las imágenes de CT e incluye información clínica de los pacientes como edad, historial de tabaquismo y antecedentes de cáncer<sup>30</sup>. Además, Sardar Hamidian *et al.* desarrollaron un modelo que combina la segmentación y clasificación de nódulos pulmonares usando una red neuronal convolucional 3D para la detección automática de

---

<sup>28</sup> Penghua ZHAI et al. "Multi-task learning for lung nodule classification on chest CT". in: *IEEE access* 8 (2020), pp. 180317–180327.

<sup>29</sup> Asghar Ali SHAH et al. "Deep learning ensemble 2D CNN approach towards the detection of lung cancer". In: *Scientific Reports* 13.1 (2023), p. 2987.

<sup>30</sup> Chao TONG et al. "Pulmonary nodule classification based on heterogeneous features learning". In: *IEEE Journal on Selected Areas in Communications* 39.2 (2020), pp. 574–581.

nódulos pulmonares en tomografías computarizadas del tórax. Este modelo aprende y captura detalles estructurales de los nódulos desde múltiples ángulos, identificando patrones de forma y volumen<sup>31</sup>. No obstante, aunque capturan patrones estructurales, no proporcionan medidas geométricas del nódulo pulmonar; e incluso, los operadores convolucionales tradicionales se enfocan en detectar patrones locales, omitiendo características de larga dependencia que podrían mejorar la clasificación y segmentación de los nódulos pulmonares.

Es por esto, que han emergido metodologías basadas en atención, como por ejemplo, la propuesta por Fu *et al.* que se enfocan en atributos específicos de la morfología tales como la lobulación, espiculación, estructura interna, entre otros para un análisis detallado de la forma del nódulo<sup>9</sup>. Además, Zia UrRehman *et al.* desarrollaron una CNN profunda con mecanismos duales de atención, combinando atención de canal y atención espacial, para extraer con precisión la forma de los nódulos pulmonares en imágenes de TC. La atención de canal se enfoca en identificar las características más relevantes de la imagen, como contornos y texturas, que son fundamentales para definir la forma de los nódulos. Por otro lado, la atención espacial dirige la atención del modelo hacia las regiones específicas donde estas características de forma son más prominentes, mejorando la capacidad del modelo para discernir detalles en la forma de los nódulos<sup>32</sup>. No obstante, estos mecanismos se rigen por esquemas supervisados que pueden ser sesgados por la información de malignidad en las etiquetas.

Otras metodologías han incluido la segmentación de nódulos y características radiológ-

---

<sup>31</sup> Sardar HAMIDIAN et al. "3D convolutional neural network for automatic detection of lung nodules in chest CT". in: *Medical Imaging 2017: Computer-Aided Diagnosis*. Vol. 10134. SPIE. 2017, pp. 54–59.

<sup>32</sup> Zia URREHMAN et al. "Effective lung nodule detection using deep CNN with dual attention mechanisms". In: *Scientific Reports* 14.1 (2024), p. 3934.

icas para aproximar la clasificación de nódulos. Por ejemplo, Jingqi Song *et al.* incluyó segmentaciones del nódulo, así como características de forma y textura, como la varianza de gris, la entropía del histograma y la circularidad, para construir un vector de características y realizar una clasificación binaria <sup>33</sup>. También se han propuesto un modelo que combina la segmentación del nódulo, con descriptores de ingeniería de características como la Matriz de Co-ocurrencia de Niveles de Gris (GLCM) y los Patrones Binarios Locales (LBP), para luego hacer una clasificación con una máquina de soporte vectorial <sup>34</sup>. En otro trabajo, se propuso la combinación de características de textura y forma, extraídas de imágenes de PET/CT para distinguir entre nódulos benignos y malignos <sup>35</sup>. De igual manera, se han desarrollado metodologías utilizando características radiómicas para la clasificación de nódulos pulmonares. Por ejemplo, Qing-qing Xu *et al.* implementaron modelos de Regresión Logística y SVM para clasificar nódulos pulmonares en benignos y malignos, basándose en la extracción de características como textura, forma, y otras métricas adicionales como la lobulación de imágenes de CT <sup>10</sup>. Estos enfoques han demostrado la relevancia de incluir información relacionada con la forma y delineación de los nódulos. Sin embargo, estos enfoques integran información de la malignidad, o información de otras modalidades, lo que puede evitar la generalización de los enfoques y el aprovechamiento de la información geométrica.

---

<sup>33</sup> Jingqi SONG et al. "Weakly-supervised classification of pulmonary nodules based on shape characters". In: *2016 IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*. IEEE. 2016, pp. 228–232.

<sup>34</sup> Muhammad Attique KHAN et al. "VGG19 network assisted joint segmentation and classification of lung nodules in CT images". In: *Diagnostics* 11.12 (2021), p. 2208.

<sup>35</sup> Yavuz Sami SALIHOĞLU et al. "Diagnostic performance of machine learning models based on 18F-FDG PET/CT radiomic features in the classification of solitary pulmonary nodules". In: *Molecular Imaging and Radionuclide Therapy* 31.2 (2022), p. 82.

## 2. PROBLEMA DE INVESTIGACIÓN

El cáncer de pulmón es la principal causa de muerte por cáncer a nivel mundial<sup>1</sup>. Uno de los principales biomarcadores para la localización, estratificación y seguimiento del cáncer de pulmón son los nódulos pulmonares observados en estudios de CT. Estos nódulos de acuerdo a la morfología, la densidad y el tamaño pueden asociarse con la malignidad del biomarcador. Sin embargo, la interpretación se realiza típicamente de forma observacional, reportando una alta variabilidad inter-intra observador (con un coeficiente kappa promedio de 0.51 y 0.57 respectivamente) al momento de anotar o estratificar un nódulo pulmonar, lo que conlleva a la posibilidad de diagnósticos incorrectos.

Recientemente, se han desarrollado metodologías computacionales basadas en representaciones convolucionales las cuales han demostrado ser eficaces para clasificar la malignidad de nódulos pulmonares bajo enfoques supervisados. Sin embargo, estos métodos pueden aprender sesgos asociados a las anotaciones. Además, estos enfoques no aprovechan de forma explícita marcadores radiológicos de los nódulos, como las características de forma. De hecho, estas arquitecturas convergen en descriptores codificados para la clasificación, pero no para el modelamiento de forma. Lo anterior puede conducir a esquemas sobreaprendidos sin un carácter de generalización.

De esta manera, surge la siguiente pregunta de investigación:

¿Cómo implementar una estrategia computacional para clasificar nódulos pulmonares entre malignos y benignos observados en tomografías computarizadas usando descriptores embebidos extraídos de arquitecturas dedicadas a la segmentación?

### **3. OBJETIVOS**

#### **3.1. Objetivo general**

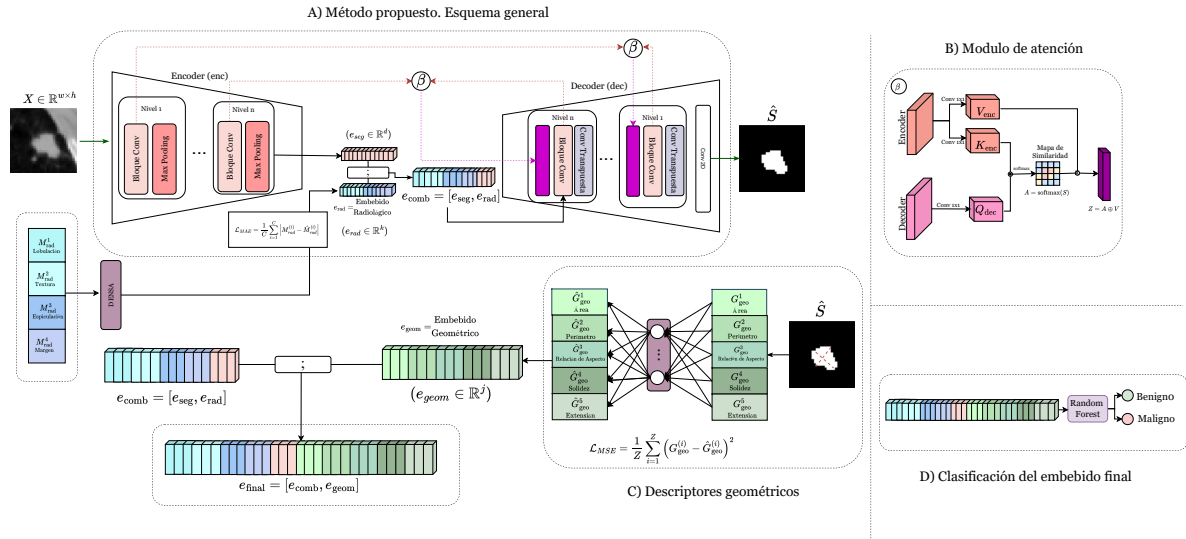
Desarrollar una estrategia para discriminar nódulos pulmonares entre malignos y benignos entrenando una arquitectura profunda dedicada a la segmentación de masas observadas en tomografías computarizadas.

#### **3.2. Objetivos específicos**

- Seleccionar un conjunto de datos públicos anonimizado de estudios de CT que incluya la localización, segmentación y estratificación de los nódulos observados.
- Desarrollar un modelo dedicado a la segmentación de nódulos pulmonares para la codificación de patrones texturales y de forma.
- Codificar representaciones embebidas con información textural y de forma para la clasificación de la malignidad de los nódulos pulmonares.
- Validar la metodología desarrollada en cuanto a la capacidad de clasificar la malignidad de los nódulos desde los descriptores de forma.

## 4. MÉTODO PROPUESTO

En el presente trabajo se implementó una estrategia de clasificación de nódulos pulmonares que utiliza descriptores de forma, compilados desde una arquitectura dedicada a la segmentación de nódulos. La arquitectura dedicada a la segmentación se basa en un esquema de autoencoder, integrando conexiones residuales (como en la U-Net) y estableciendo mecanismos de atención entre el codificador y decodificador. Además, el aprendizaje de esta red es reforzado por *tokens* radiológicos que guían la segmentación y permiten integrar observaciones semánticas sobre la forma de los nódulos. Desde los embebidos de esta arquitectura y junto con medidas tomadas a la delineación resultante, se conforma el descriptor de forma, que posteriormente será mapeado a un bosque aleatorio (*Random forest*) para determinar la malignidad del nódulo. La metodología propuesta se ilustra en la Figura 3



**Figura 3.** Esquema general de la estrategia implementada, donde se integran descriptores aprendidos y anotaciones radiológicas. B) Detalle del módulo de atención implementado  $\beta$ . C) Extracción de los descriptores geométricos calculados a partir de las segmentaciones generadas por el modelo. D) Validación mediante un clasificador Random Forest.

#### 4.1. Autoencoder para determinación de forma nodular

La principal hipótesis de este trabajo yace en que una apropiada caracterización de la forma del nódulo puede brindar elementos necesarios para codificar un descriptor que tenga una correlación asociada con la malignidad. Siguiendo este punto, en este trabajo, la caracterización de la forma de los nódulos pulmonares se realizó utilizando una arquitectura U-Net. Esta red fue seleccionada por su capacidad para preservar detalles espaciales y capturar patrones texturales relevantes, a través de su arquitectura de codificador-decodificador, reforzada por conexiones residuales. En cuanto al codificador, la entrada al modelo, denotada como  $\mathbf{X} \in \mathbb{R}^{H \times W}$ , representa un parche bidimensional del nódulo, donde  $H$  y  $W$  son las dimensiones de la imagen. En el transcurso del proceso de codificación, cada capa  $l \in \{1, \dots, L\}$  transforma la entrada en un conjunto de características  $\Phi_k^\ell$  mediante operaciones convolucionales parametrizadas por un filtro  $\Psi_k^\ell$  y un sesgo  $b_k^\ell$ . Estas operaciones, combinadas con activaciones no lineales ( $a_k^\ell$ ) y operadores de *pooling* ( $\Pi_\ell$ ), generan características intermedias que encapsulan información espacial y textural del nódulo. Formalmente, el proceso en cada capa puede describirse como:  $\Phi_k^\ell = \Pi_\ell (a_k^\ell (\Phi^{\ell-1} * \Psi_{k_{enc}}^\ell + b_k^\ell))$ , donde  $\Phi^\ell$  representa las características aprendidas en la capa  $\ell$ , y  $\Phi^0 = \mathbf{X}$ .

Una vez realizado el procesamiento de las capas, el embebido final del codificador,  $e_{seg}$ , se define como  $e_{seg} = \Phi(\Phi_1 \circ \Phi_2 \circ \dots \circ \Phi_L)$ , donde  $\Phi$  representa la función compuesta que integra todas las transformaciones aplicadas por el codificador. Este embebido  $e_{seg} \in \mathbb{R}^d$  tiene una dimensión  $d$ , siendo un descriptor compacto de las características del nódulo. Este embebido actúa como una representación densa y compacta de la forma y textura del nódulo, sirviendo como base para las tareas de segmentación del nódulo. De hecho, desde esta representación se podrá asociar el descriptor de forma compacta, que luego se complementará con otra información de forma para evaluar su capacidad de clasificación de la malignidad.

Por otro lado, en la etapa de decodificación, el modelo reconstruye la segmentación final  $\hat{S}$  a partir del embebido generado por el codificador  $e_{seg}$ . Este proceso utiliza operaciones de convoluciones transpuestas que permiten recuperar progresivamente las características espaciales y texturales del nódulo pulmonar. La salida final se obtiene aplicando una función de activación sigmoide, denotada como  $g$ , de manera que  $\hat{S} = g(\Psi_{dec} * e_{seg} + b_{dec})$ , donde  $\Psi_{dec}$  y  $b_{dec}$  son los pesos y sesgos del decodificador. Estas transformaciones logran la recuperación de los detalles anatómicos necesarios para realizar una segmentación. No obstante, puede existir una pérdida de detalles de los nódulos, como su forma, debido a la concatenación cruda entre características del codificador y decodificador durante las conexiones de salto *skip-connections*. De hecho, realizar una concatenación directa entre las convoluciones de salida  $\Phi^\ell$  de cada capa puede involucrar información de fondo que no es relevante o propagar errores en la representación de la forma. Lo anterior puede estar asociado al mapeo y concatenación de mapas de representación cruda, donde se ponderan todos los píxeles por igual.

Por las razones anteriormente expuestas, se implementó un mecanismo de atención cruzada  $\beta$ , que permite que el decodificador module la importancia de las características recibidas del codificador, enfocándose en aquellas más relevantes para la segmentación. Este mecanismo de atención opera mediante tres componentes principales: consulta ( $Q_{dec}$ ), la cual se obtiene del decodificador, ya que representa la información actual de la reconstrucción en cada capa, permitiendo que la red enfoque su atención en regiones específicas de la segmentación. Mientras tanto, la clave ( $K_{enc}$ ) y valor ( $V_{enc}$ ) provienen del codificador y contienen la información estructural extraída en las primeras etapas de la red. La atención cruzada entre el codificador y decodificador se calcula como:

$$A = \text{softmax} \left( \frac{Q_{dec} K_{enc}^\top}{\sqrt{d_k}} \right) V_{enc},$$

Donde  $d_k$  es la dimensionalidad de las características. Este esquema permite capturar relaciones espaciales y estructurales entre las capas del codificador y decodificador, mejorando la representación del nódulo y la precisión de la segmentación resultante.

Para reforzar esta representación y lograr detalles locales, altamente relevantes en la caracterización de nódulos, en este trabajo se incluyó la información clínica de la valoración radiológica brindada por un experto. De esta forma, nosotros aportamos en la integración de información clínica, estrictamente relacionada con la forma, como por ejemplo el carácter lobulado o espiculado de una masa observada. Estas anotaciones incluyen información morfológica relevante. En este trabajo, estas características radiológicas se representan como un conjunto de variables  $\{M_{rad}^i\}_{i=1}^C$ , donde  $M_{rad}^i$  denota cada variable anotada y  $C$  es el número total de categorías evaluadas.

Las características radiológicas se agrupan en un vector  $[M_{rad}^1; M_{rad}^2; \dots; M_{rad}^C]$ , que se procesa mediante una capa densa para generar un embebido radiológico  $e_{rad} \in \mathbb{R}^k$ . Este embebido condensa la información morfológica en un espacio de representación continuo y de menor dimensionalidad, facilitando su integración con el embebido generado por el codificador de la U-Net ( $e_{seg}$ ). Durante el entrenamiento, ambos embebidos se concatenan para formar una representación combinada  $e_{comb} = [e_{seg}; e_{rad}]$ , permitiendo aprovechar tanto las características estructurales aprendidas automáticamente como las anotaciones expertas de los radiólogos. Esta representación combinada fortalece la capacidad del modelo para realizar una segmentación y clasificación confiables de los nódulos pulmonares, al combinar información complementaria que captura tanto patrones morfológicos como texturales.

En general, el proceso de segmentación es minimizado mediante una función de pérdida de entropía cruzada, definida como  $L_{seg} = \frac{1}{N} \sum_{i=1}^N (S_i \log(\hat{S}_i) + (1 - S_i) \log(1 - \hat{S}_i))$ . De forma complementaria, para garantizar que el embebido radiológico capture con precisión

la información proporcionada por los expertos, su proceso de aprendizaje es guiado por la función de pérdida Mean Absolute Error (MAE), definida como:  $L_{rad} = \frac{1}{C} \sum_{i=1}^C |M_{rad}^{(i)} - \hat{M}_{rad}^{(i)}|$ , donde  $M_{rad}^{(i)}$  representa las características radiológicas reales y  $\hat{M}_{rad}^{(i)}$  son las predicciones del modelo. Así, la función de pérdida total es definida como  $L_T = L_{rad} + L_{seg}$ . Esta función de pérdida ayuda a preservar la información morfológica en el embebido, evitando la influencia de errores muy grandes durante la optimización.

## 4.2. Descriptores geométricos

De forma complementaria, en este trabajo se modeló la información geométrica relevante a partir de las segmentaciones obtenidas, sin requerir un entrenamiento explícito basado en etiquetas de malignidad. De esta forma se busca codificar descriptores robustos al sesgo que pueden causar las anotaciones brindadas por los radiólogos. Así, la estrategia propuesta comienza con la segmentación de los nódulos  $\hat{S}$ , en donde se extraen diversas medidas geométricas que describen la estructura de los nódulos pulmonares, proporcionando una representación complementaria a la obtenida mediante codificación profunda.

Estas características geométricas son definidas como:  $[G_{geo}^1, G_{geo}^2, \dots, G_{geo}^Z]$ , donde  $Z$  representa el número de características calculadas. Desde este vector de medidas geométricas se realiza un aprendizaje auto-supervisado, que permite aprender un embebido compacto con la información geométrica. Teniendo en cuenta la dimensionalidad de las características geométricas, se decidió aprender una única capa oculta embebida, desde donde se genera el vector geométrico  $e_{geom} \in \mathbb{R}^j$  capturando las características de forma relevantes de los nódulos pulmonares. Para garantizar que este embebido geométrico preserve la información estructural contenida en las segmentaciones, su proceso de aprendizaje se optimiza mediante la función de pérdida de Error Cuadrático Medio (MSE), definida como:  $\mathcal{L}_{MSE} = \frac{1}{Z} \sum_{i=1}^Z (G_{geo}^{(i)} - \hat{G}_{geo}^{(i)})^2$ , donde  $G_{geo}^{(i)}$  representa los valores reales de las características geométricas y  $\hat{G}_{geo}^{(i)}$  son las predicciones del modelo. El autoencoder

para las características geométricas, y su ilustración con la función de pérdida, puede observarse en la Figura 3

### 4.3. Clasificación de descriptor de forma

El descriptor propuesto entonces en este trabajo consiste en la concatenación de los embebidos calculados previamente, es decir, el embebido que resulta de la tarea de segmentación, junto con la información morfológica  $e_{\text{comb}}$  y el embebido calculado desde la información geométrica  $e_{\text{geom}}$ . Entonces, la concatenación resultante se expresa como:  $e_{\text{final}} = [e_{\text{seg}}, e_{\text{rad}}, e_{\text{geom}}]$ , asegurando que la representación final encapsule atributos estructurales, geométricos y morfológicos en un espacio de representación compacto y enriquecido de los nódulos pulmonares.

Entonces, con el fin de validar la capacidad de discriminación del descriptor propuesto ( $e_{\text{final}}$ ), en cuanto a la separación de nódulos malignos, se realizó una proyección de estos descriptores a un algoritmo clásico de aprendizaje de máquina. En este caso, se seleccionó un bosque aleatorio (RF, por sus siglas en inglés de random forest). Esta técnica es ampliamente utilizada en el ámbito médico debido a su robustez y eficacia en tareas de clasificación, permitiendo manejar datos multi-dimensionales, heterogéneos de forma ensamblada. Además, su construcción basada en múltiples árboles de decisión entrenados sobre diferentes subconjuntos de datos lo hace menos propenso al sobre-ajuste. También, su carácter de explicación permite brindar un mayor soporte y adopción en la rutina clínica. Particularmente, en este trabajo se entrenó el bosque aleatorio con tuplas supervisadas  $(e_{\text{final}}^k, y^i)$ , donde  $y^i$  son etiquetas (benignos, malignos). En este caso, se realiza la generación de múltiples árboles de decisión, el cual es definido matemáticamente como  $\{h_t(\cdot)\}_{t=1}^T$ , donde cada árbol representa una posible estructura de decisión basada en las características de los nódulos. Cada árbol de decisión no se entrena con la

totalidad de los datos disponibles, sino con un subconjunto creado mediante el muestreo *bootstrap*. Este método consiste en tomar muestras aleatorias del conjunto de datos original con reemplazo, lo que significa que algunas muestras pueden aparecer más de una vez en el mismo subconjunto. De esta manera, cada árbol aprende con datos ligeramente diferentes, lo que introduce diversidad en el modelo permitiendo que se aprendan patrones complejos de los nódulos pulmonares.

En cada nodo del árbol de decisión, que representa un punto donde se evalúa una característica específica del nódulo pulmonar, se selecciona aleatoriamente un subconjunto de características  $k_i$  del embebido final  $e_{\text{final}}$ . La cantidad de características en este subconjunto está determinada por el hiperparámetro  $\theta$ , el cual define cuántos de estos atributos se consideran al momento de realizar una división. Esta selección aleatoria reduce la correlación entre los árboles, mejorando la capacidad de generalización del modelo. Como función de optimización se usó la Ganancia de información  $= H(S) - \sum_{j=1}^m \frac{|S_j|}{|S|} H(S_j)$ , donde  $S_j$  representa los subconjuntos generados tras la partición y  $H(S_j)$  su entropía. En este caso, la ganancia de información permite identificar qué características del embebido final  $e_{\text{final}}$  contribuyen más a la separación de los nódulos en el espacio de decisión. Formalmente, la predicción final del RF se obtiene mediante la agregación de las predicciones individuales de los árboles de decisión utilizando un voto mayoritario, definido como  $\hat{y} = \text{mode}\{h_t(e_{\text{final}}), t = 1, \dots, T\}$ , donde  $h_t(e_{\text{final}})$  representa la clasificación de cada árbol sobre el embebido final de los nódulos pulmonares. Además, el RF facilita la interpretación del modelo al calcular la importancia relativa de cada característica en el proceso de decisión, resultando fundamental para comprender qué factores desempeñan un papel clave en la discriminación de los patrones aprendidos por la red.

#### 4.4. Diseño Experimental

**Datos.** En este trabajo se consideró el conjunto de datos público LIDC-IDRI (Lung Image Database Consortium Image Collection), que consta de 1018 scans de tomografía computarizada (CT) de baja dosis correspondiente a 1010 pacientes <sup>36</sup>. En este conjunto de datos, un total de cuatro radiólogos expertos realizaron la anotación y caracterización de la localización, diagnóstico y forma (lobulación, espiculación, textura, entre otros) de los nódulos pulmonares. Así, el dataset es consistente en cuanto al acuerdo de las observaciones radiológicas y también considera información extra sobre la valoración de cada estudio, permitiendo complementar el estudio. Este conjunto de datos incluye un total de 2612 nódulos pulmonares anotados, de los cuales 887 cuentan con anotaciones realizadas por los cuatro radiólogos involucrados en el estudio. Cada nódulo fue evaluado según una escala de malignidad de 1 a 5, donde el 8% corresponde a nódulos altamente benignos (1), el 12% a moderadamente benignos (2), el 42.39% a nódulos indeterminados (3), el 17% a moderadamente malignos (4), y el 23% a altamente malignos (5). Además, se incluyeron anotaciones sobre características morfológicas relevantes para el diagnóstico, como lobulación, espiculación, textura, margen, calcificación y sutileza, las cuales también están caracterizadas en una escala entre el uno y el cinco.

Para garantizar la consistencia en las etiquetas de malignidad, se utilizó el criterio Median Malignancy Value (MMV), que se basa en calcular el valor mediano de las puntuaciones

---

<sup>36</sup> Samuel G ARMATO III et al. "The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans". In: *Medical physics* 38.2 (2011), pp. 915–931.

otorgadas por al menos tres radiólogos<sup>37 38</sup>. Este enfoque es ampliamente utilizado en la literatura para tener un mayor valor de certeza en cuanto a la valoración subjetiva realizada por los radiólogos. En este trabajo, se decidió incluir únicamente los nódulos pertenecientes a las categorías 1-2 (benignos) y 4-5 (malignos), excluyendo los nódulos clasificados como indeterminados (clase 3). Esta exclusión se debe a la ambigüedad inherente de la clase, ya que presentan características mixtas que dificultan su clasificación, lo que podría introducir ruido en el modelo y afectar su rendimiento. Además, se seleccionaron las características lobulación, espiculación, margen y textura, ya que describen la morfología y contornos del nódulo. En contraste, la calcificación fue excluida porque está relacionada con la composición interna del nódulo más que con su forma y, además, en el conjunto de datos, la mayoría de los nódulos fueron etiquetados con el nivel 6 (ausencia de calcificación), lo que introduce un sesgo en su distribución. La sutileza también se descartó, ya que mide la facilidad con la que un radiólogo puede identificar un nódulo en la imagen, lo cual depende más del contraste y la calidad de la imagen que de su estructura. Esta selección optimiza el enfoque del modelo al priorizar características con mayor valor discriminativo en la tarea propuesta.

**Preprocesamiento** Para garantizar la estandarización en el procesamiento de los datos, las imágenes y las máscaras de los nódulos pulmonares fueron preprocesadas, siguiendo criterios estándar en el análisis de estos tipos de imágenes y el dominio de nódulos pulmonares. Este proceso incluyó la selección de cortes específicos, el ajuste de di-

---

<sup>37</sup> Raunak DEY; Zhongjie LU, and Yi HONG. “Diagnostic classification of lung nodules using 3D neural networks”. In: *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*. IEEE. 2018, pp. 774–778.

<sup>38</sup> Raul Victor M DA NOBREGA et al. “RETRACTED ARTICLE: Lung nodule malignancy classification in chest computed tomography images using transfer learning and convolutional neural networks”. In: *Neural Computing and Applications* 32.15 (2020), pp. 11065–11082.

mensiones de entrada y la normalización de las características extraídas. Las imágenes utilizadas en este estudio corresponden a cortes axiales  $2D$  de CT, en las que se identifican regiones de interés centradas en los nódulos pulmonares. Con el fin de capturar la información más representativa del nódulo, se seleccionó el corte medio dentro de cada conjunto de imágenes. La elección de este corte se basó en la premisa de que el centro del conjunto de cortes contiene la mayor extensión del nódulo, lo que permite extraer una mejor caracterización de su estructura. Por último, los parches extraídos presentan una resolución de  $32 \times 32$  píxeles para mantener una representación uniforme y permitir un procesamiento eficiente en la arquitectura propuesta.

Para garantizar una escala homogénea entre las diferentes entradas al modelo, se aplicaron las siguientes técnicas de normalización adecuadas para cada tipo de dato:

1. **Parches de imagen:** Se utilizó una normalización min-max, escalando los valores de intensidad de los píxeles al rango  $[0, 1]$ , definida como:  $X_{img} = \frac{X - X_{min}}{X_{max} - X_{min}}$
2. **Descriptores radiológicos:** Los valores asociados a la severidad de las características radiológicas (lobulación, textura, espiculación y margen), fueron normalizadas mediante estandarización  $X_{rad} = \frac{X - \mu}{\sigma}$ . Donde  $X$  representa el valor de la característica,  $\mu$  la media y  $\sigma$  la desviación estándar calculada sobre el conjunto de datos permitiendo obtener una representación homogénea entre las características discretas.

**Configuración de la arquitectura propuesta** Para cada rama de la arquitectura propuesta, se realizó la siguiente configuración:

1. **Segmentación morfológica:** El codificador está compuesto por cuatro niveles, cada uno con dos capas convolucionales de tamaño de *kernel*  $3 \times 3$ , activación *ReLU* y normalización por *batches*. La reducción de la resolución espacial se realizó mediante *Max Pooling* de  $2 \times 2$  en cada nivel, permitiendo una extracción jerárquica de

características a diferentes escalas. Al finalizar la fase de codificación, se obtiene el embebido de segmentación  $e_{\text{seg}} \in \mathbb{R}^{2 \times 2 \times 96}$ .

En cuanto al decodificador, este componente sigue una estructura simétrica, aplicando operaciones de *upsampling* mediante convoluciones transpuestas de tamaño  $3 \times 3$ . Cada nivel del decodificador está compuesto por dos capas convolucionales con activación *ReLU* y normalización por *batches*. Adicionalmente, en cada nivel se emplearon conexiones tipo *skip-connection*, que transfieren características de alta resolución desde el codificador al decodificador, preservando detalles espaciales en la segmentación final. La segmentación final se obtiene mediante una convolución  $1 \times 1$  con activación *sigmoide*.

2. **Integración entre codificador-decodificador:** Para mejorar la segmentación de los nódulos pulmonares, se incorporaron mecanismos de atención cruzada en la arquitectura propuesta. Se implementaron un total de cuatro bloques de atención cruzada, uno por cada nivel del decodificador, donde en cada bloque se calcularon los tensores  $k$  (clave),  $q$  (consulta) y  $v$  (valor). La clave y el valor ( $k, v$ ) se obtienen de la representación del codificador en el mismo nivel, mientras que la consulta ( $q$ ) proviene del decodificador; cada tensor fue proyectado mediante convoluciones  $1 \times 1$ .
3. **Embebido morfológico** Adicionalmente, la arquitectura incorpora descriptores radiológicos para mejorar la caracterización de los nódulos pulmonares. Los descriptores radiológicos son procesados mediante una capa densa con 56 neuronas y activación *ReLU*, generando un embebido radiológico  $e_{\text{rad}} \in \mathbb{R}^{2 \times 2 \times 56}$ . Este embebido se optimiza mediante la función de pérdida MAE.
4. **Embebido geométrico:** A partir de las segmentaciones generadas por la arquitectura propuesta, se extraen descriptores geométricos que permiten modelar la forma

del nódulo sin requerimientos de etiquetas de malignidad. En este caso, se calculan las siguientes características:

- (a) El área ( $A$ ) representa la cantidad total de píxeles contenidos dentro del contorno del nódulo segmentado, definida como  $A(\hat{S}) = \sum_{(i,j) \in C(\hat{S})} 1$ , donde  $C(\hat{S})$  es el conjunto de píxeles que conforman el contorno del nódulo. Este descriptor permite cuantificar el tamaño del nódulo en función de los píxeles segmentados.
- (b) El perímetro ( $P$ ) se obtiene sumando las distancias entre los puntos consecutivos que conforman el contorno del nódulo, lo que permite describir la longitud de su borde externo. Se define como  $P(\hat{S}) = \sum_{k=1}^N \|(x_k, y_k) - (x_{k+1}, y_{k+1})\|$ , donde  $(x_k, y_k)$  y  $(x_{k+1}, y_{k+1})$  son puntos consecutivos en el contorno del nódulo y  $N$  es el número total de puntos en dicho contorno. Un perímetro más alto puede indicar una forma más irregular o compleja.
- (c) La relación de aspecto ( $AR$ ) mide la proporción entre el ancho y la altura del nódulo, definida como  $AR(\hat{S}) = \frac{w}{h}$ , donde  $w$  y  $h$  representan el ancho y la altura del rectángulo delimitador del nódulo. Un valor cercano a 1 sugiere una forma proporcional (aproximando un círculo), mientras que valores más altos o más bajos indican estructuras alargadas en diferentes direcciones permitiendo discernir entre benigno y maligno.
- (d) La solidez ( $S$ ) cuantifica qué tan compacta es la forma del nódulo en relación con su envolvente convexa, definida como  $S(\hat{S}) = \frac{A(\hat{S})}{A_{\text{convexo}}(\hat{S})}$ , donde  $A(\hat{S})$  es el área del nódulo y  $A_{\text{convexo}}(\hat{S})$  es el área de su envolvente convexa. Valores cercanos a 1 indican que el nódulo es compacto y no presenta irregularidades significativas, mientras que valores más bajos sugieren la presencia de bordes irregulares o protuberancias.
- (e) La extensión ( $E$ ) mide la relación entre el área del nódulo y el área total del rec-

tángulo delimitador mínimo, definida como  $E(\hat{S}) = \frac{A(\hat{S})}{w \times h}$ , donde  $A(\hat{S})$  representa el área del nódulo y  $w \times h$  es el área total del rectángulo delimitador.

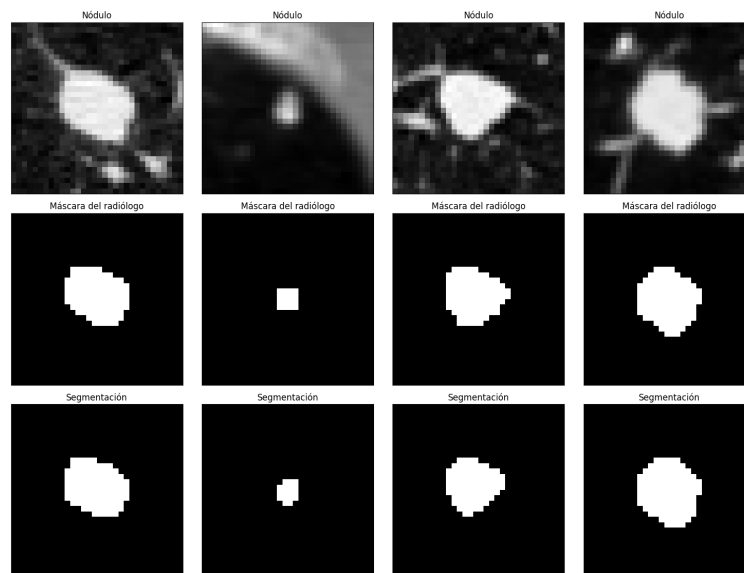
Estos descriptores geométricos se normalizan y se procesan mediante una capa densa de 28 neuronas, generando un embebido geométrico  $e_{\text{geom}} \in \mathbb{R}^{28}$ . Este embebido es optimizado mediante la función de pérdida MSE.

**Entrenamiento y validación de la arquitectura** El entrenamiento de la arquitectura se realizó utilizando el optimizador Adam con una tasa de aprendizaje inicial de 0.0001, que se mantuvo constante durante las primeras 50 épocas. Para evaluar de forma robusta y prevenir el sobre-ajuste, se empleó un esquema de validación *k-fold* con 10 divisiones, realizando un total de 100 épocas por cada pliegue. Además, se implementó una estrategia de *early-stopping* después de 10 épocas consecutivas. Adicionalmente, se realizó un *data augmentation* incluyendo transformaciones geométricas en donde se realizaron rotaciones verticales, horizontales y en diferentes ángulos [90, 180, 270].

Adicionalmente, se utilizó un *Random Forest* para la validación del modelo, configurado con los parámetros predeterminados, incluyendo un número de 200 árboles de decisión, con una profundidad de 14.44 y una selección aleatoria de características para cada árbol ( $max\_features = 'sqrt'$ ). Para esta evaluación se utilizaron las mismas particiones del *k-fold*, garantizando una adecuada partición de los datos a lo largo de la aplicación de la metodología. Las métricas utilizadas para evaluar el desempeño del Random Forest incluyeron AUC (Área bajo la curva ROC), precisión, sensibilidad y F1 Score. Estas métricas se calcularon en cada iteración de la validación cruzada, registrando la media y la desviación estándar de los resultados para reflejar la estabilidad y la robustez del modelo en diferentes particiones del conjunto de datos.

## 5. EVALUACIÓN Y RESULTADOS

En este trabajo se propuso una metodología para construir descriptores que discriminan la malignidad del nódulo, pero cuyo entrenamiento estuvo explícitamente basado en la segmentación y caracterización de la forma de estas masas observadas. Además, la metodología incluyó como refuerzo la integración de información radiológica que los expertos brindan sobre los hallazgos morfológicos de los nódulos. En primera instancia, en la evaluación del método se validó la capacidad de recuperar la segmentación de los nódulos pulmonares. En la Figura 4 se ilustran ejemplos representativos de segmentaciones generadas por el modelo frente a las anotaciones manuales. Como se observa, el método logra codificar incluso características locales de la geometría del nódulo, lo cual puede ser relevante a la hora de categorizar su malignidad. No obstante, en algunas situaciones se aprecian discrepancias sutiles en los márgenes de la segmentación predicha con respecto a la máscara de referencia.



**Figura 4.** Ejemplos de segmentaciones predichas por la U-Net en diferentes casos de nódulos pulmonares.

En términos cuantitativos, en la tabla 2 se ilustran los resultados obtenidos en términos de la capacidad de segmentación del método propuesto. Además, se hace una comparación de este método con respecto a una U-net de línea base, así como también el esquema de atención sin la inclusión de los hallazgos radiológicos. Como se observa, el método propuesto alcanza un Dice Score (DsC) de 95.80%, mostrando coherencia en los nódulos segmentados con respecto a las marcaciones realizadas por un experto. También se logra evidenciar que la información de los hallazgos radiológicos, así como las conexiones de atención entre el codificador y el decodificador, brindan una mayor eficacia a la hora de obtener las segmentaciones propuestas. Teniendo en cuenta que en nuestro trabajo se toman medidas geométricas desde las segmentaciones, y que esta información acompaña a los embebidos obtenidos de la imagen, es de vital importancia tener una buena representación, para así lograr una discriminación de malignidad sin el uso explícito de estas etiquetas.

<b>Modelo</b>	<b>Dice Score (%)</b>	<b>Hausdorff (px)</b>
U-Net línea base	91.33 ± 0.009	1.58 ± 0.120
U-Net con mecanismo de atención	93.21 ± 0.016	1.32 ± 0.168
<b>Método propuesto</b>	<b>95.80 ± 0.006</b>	<b>1.10 ± 0.087</b>

**Tabla 2.** Resultados del Dice Score y Hausdorff para las diferentes configuraciones de la U-Net.

En cuanto a la evaluación de discriminación de la malignidad, como objetivo principal en este trabajo, se decidió realizar inicialmente un análisis exhaustivo de la capacidad de los descriptores de forma de separar estas clases. Como base de conocimiento, inicialmente se obtuvieron medidas geométricas sobre las anotaciones realizadas por un experto, las cuales luego fueron mapeadas a un Random Forest (primera línea en la fila) . También se clasificaron los hallazgos radiológicos brindados por expertos durante la inspección de rutina de un estudio CT (segunda fila en la tabla). Con estas dos bases, se puede observar que la información morfológica y geométrica de los nódulos resulta relevante para discriminar la malignidad del nódulo, lo cual hace que resulte relevante codificar

arquitecturas que se dediquen a esta caracterización, evitando así el sesgo brindado por las anotaciones sobre malignidad realizadas por expertos. Entonces, en la tabla 3 se muestran los resultados del método propuesto, que con una ganancia del 10.22% de AUC y de 8.1% en F1-score, muestra ser competitiva para reunir información morfológica, geométrica y codificar un descriptor eficaz en la tarea de clasificación de la malignidad.

	<b>AUC (%)</b>	<b>Especificidad (%)</b>	<b>Sensibilidad (%)</b>	<b>Precisión (%)</b>	<b>F1-Score (%)</b>
Características geométricas	82.40 ± 0.022	83.22 ± 0.032	89.27 ± 0.036	84.22 ± 0.015	86.61 ± 0.014
Hallazgos radiológicos	83.23 ± 0.031	87.81 ± 0.026	85.80 ± 0.050	90.38 ± 0.056	87.83 ± 0.034
<b>Método propuesto</b>	<b>93.02 ± 0.005</b>	<b>91.18 ± 0.001</b>	94.86 ± 0.010	<b>95.79 ± 0.005</b>	<b>95.32 ± 0.005</b>

**Tabla 3.** Evaluación individual del desempeño de los descriptores geométricos y radiológicos utilizados como entrada al clasificador, en comparación con el método propuesto que combina ambos conjuntos de características.

En un estudio de componentes de la arquitectura propuesta, se validaron con componentes específicos del método propuesto. Como se puede ver en la Tabla 4, se analizaron cinco configuraciones diferentes para evaluar el impacto de cada sub-módulo, en cuanto a la capacidad de discriminación de nódulos pulmonares. Como se puede observar, la línea base, dedicada solo a la segmentación, no logra tener una codificación discriminativa en el embebido codificado. Lo anterior puede deberse a la alta variabilidad de los nódulos y la carencia de otras características globales de estas masas. Una vez se le adiciona al modelo los mecanismos de atención, se mejora entre las capas del encoder y decoder de la U-Net; el AUC aumenta a 70.94%, mejorando también el F1-score a 84.35%. Esta mejora indica que la atención permite capturar mejor la información contextual de los nódulos. Posteriormente, al mantener la U-Net con atención e integrar descriptores geométricos extraídos de las segmentaciones, el AUC aumenta a 79.32%, lo cual indica que esta adición enriqueció la representación del nódulo al proporcionar características complementarias. De igual manera, al incorporar anotaciones radiológicas, el modelo alcanzó un AUC de 83.87% y un F1-score de 88.80%, lo que nos indica que la integración de esta información morfológica incrementó la capacidad del modelo para discernir pa-

trones complejos en los nódulos pulmonares. En el último caso, considerando todos los componentes, el método propuesto logra los valores más altos en todas las métricas, con un AUC, F1-score y especificidad de 93.02%, 95.32% y 91.18% respectivamente, demostrando presentar una alta capacidad de discriminación. Esto, permitiendo capturar información espacial y estructural de los nódulos, complementándola con características cuantificadas de sus formas y texturas. Adicionalmente, permite obtener un modelo más robusto y denso, optimizando la clasificación.

	<b>AUC (%)</b>	<b>Especificidad (%)</b>	<b>Sensibilidad (%)</b>	<b>Precisión (%)</b>	<b>F1-Score (%)</b>
U-Net línea base	59.96 ± 0.050	56.53 ± 0.016	96.22 ± 0.033	68.95 ± 0.058	80.17 ± 0.039
U-Net con mecanismo de atención	70.94 ± 0.020	78.31 ± 0.023	88.82 ± 0.027	80.37 ± 0.017	84.35 ± 0.013
U-Net basada en geometría	79.32 ± 0.019	86.50 ± 0.005	81.89 ± 0.018	85.88 ± 0.016	83.81 ± 0.009
U-Net basada en morfología radiológica	83.87 ± 0.018	88.42 ± 0.05	88.29 ± 0.010	89.32 ± 0.014	88.80 ± 0.011
<b>Método propuesto</b>	<b>93.02 ± 0.005</b>	<b>91.18 ± 0.001</b>	<b>94.86 ± 0.010</b>	<b>95.79 ± 0.005</b>	<b>95.32 ± 0.005</b>

**Tabla 4.** Comparación del desempeño de diferentes configuraciones de U-Net: línea base, con mecanismo de atención, integrando descriptores geométricos o morfológicos radiológicos, y el método propuesto que combina ambas fuentes de información.

Paso seguido, el siguiente análisis consistió en evaluar el método propuesto con respecto a algunos enfoques recientes del estado del arte . La Tabla 5 resume los resultados de distintos trabajos representativos, incluyendo tanto metodologías supervisadas (Tang y Ma *et. al.*)<sup>39 40</sup> como semi-supervisadas (Apostolopus y Shiet. *al.*)<sup>41 42</sup> En este caso, se puede observar que el método propuesto reporta resultados competitivos en algunas

<sup>39</sup> Guozhi TANG et al. “Multi-type classification of lung nodules based on CT radiomics and ensemble learning for diversity weighting”. In: *Quantitative Imaging in Medicine and Surgery* 14.12 (2024), p. 8942.

<sup>40</sup> Ling MA et al. “A novel fusion algorithm for benign-malignant lung nodule classification on CT images”. In: *BMC Pulmonary Medicine* 23.1 (2023), p. 474.

<sup>41</sup> Ioannis D APOSTOLOPOULOS; Nikolaos D PAPATHANASIOU, and George S PANAYIOTAKIS. “Classification of lung nodule malignancy in computed tomography imaging utilising generative adversarial networks and semi-supervised transfer learning”. In: *Biocybernetics and Biomedical Engineering* 41.4 (2021), pp. 1243–1257.

<sup>42</sup> Feng SHI et al. “Semi-supervised deep transfer learning for benign-malignant diagnosis of pulmonary nodules in chest CT images”. In: *IEEE transactions on medical imaging* 41.4 (2021), pp. 771–781.

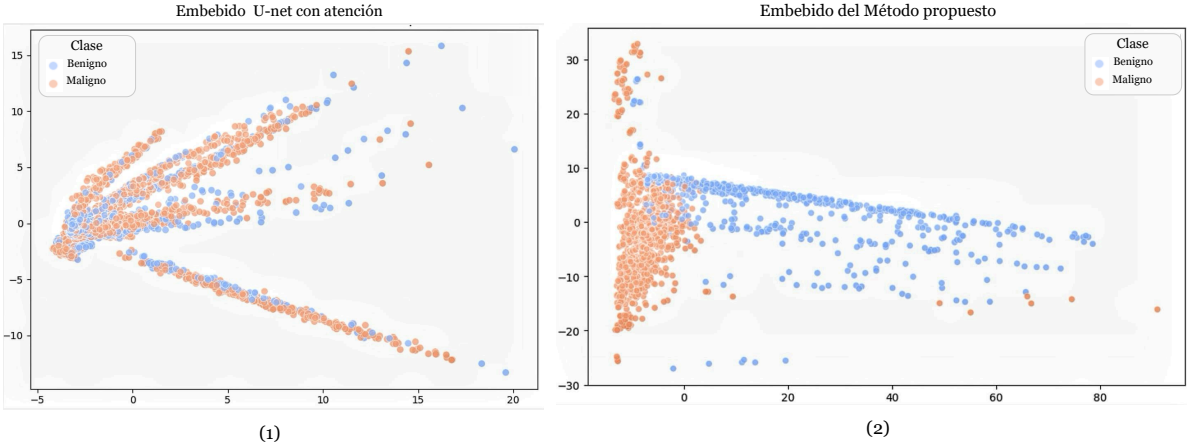
métricas, siendo favorable para el F1-score (95.32%), el Recall (94.86%) y la Precisión (95.79%). Cabe resaltar que en nuestro caso, el método propuesto no utiliza de forma explícita las etiquetas para realizar la clasificación, sino que aprende un descriptor basado en la información de forma. Con respecto a los modelos semi-supervisados, el método de Shi *et. al.* <sup>42</sup> es una arquitectura 3D que representa los nódulos en un espacio latente y utiliza grafos para conectar muestras etiquetadas y no etiquetadas según su similitud. A través de una red neuronal sobre grafos (GCN), el modelo propaga información entre estas estructuras, permitiendo que los nódulos no etiquetados se beneficien del conocimiento de aquellos con diagnóstico conocido. Con respecto a los modelos exclusivamente supervisados, las metodologías proponen un enfoque basado en características radiómicas (*radiomics*) con CNNs y GCNs capturando relaciones estructurales de los nódulos pulmonares <sup>40</sup>, y un modelo basado en radiomics y aprendizaje por ensamble con ponderación por diversidad <sup>39</sup> lo cual permite obtener una caracterización de la forma de los nódulos pulmonares. De hecho, el método propuesto gana en todas las métricas a excepción del AUC (3.27%) y la especificidad (4.64%) con respecto a Ma *et.al.*

Método	AUC (%)	Especificidad (%)	Precisión (%)	Recall (%)	F1-Score (%)
Tang et al. <sup>43</sup>	89.53	81.40	-	81.95	-
Ma et al. <sup>44</sup>	<b>96.29 ± 0.018</b>	<b>95.82 ± 0.032</b>	-	-	91.14 ± 0.029
Apostolopus et al. <sup>45</sup>	92.08	94.80	-	89.35	91.90 ± 0.015
Shi et al. <sup>46</sup>	91.0	78.6	91.9	92.2	91.9
<b>Método Propuesto</b>	<b>93.02 ± 0.005</b>	<b>91.18 ± 0.001</b>	<b>95.79 ± 0.005</b>	<b>94.86 ± 0.010</b>	<b>95.32 ± 0.005</b>

**Tabla 5.** Comparación del método propuesto con enfoques del estado del arte en la clasificación de nódulos pulmonares.

En un último análisis, se realizó una inspección observacional de los embebidos resultantes, mapeados en un espacio embebido de baja dimensionalidad. Para ello, se realizaron proyecciones bidimensionales del espacio latente con la finalidad de visualizar la distribución de las muestras de nódulos pulmonares en dos escenarios: (1) utilizando únicamente la representación latente producida por la U-Net con atención, y (2) empleando el embebido enriquecido del método propuesto. De este modo, se puede visualizar en la

Figura 5 que el método propuesto construye un embebido más robusto al combinar características aprendidas de la imagen con información radiológica y geométrica. Al comparar ambos escenarios, se evidencia que la U-Net con atención produce representaciones de nódulos benignos y malignos con mayor superposición, dificultando una separación clara entre clases. En contraste, el embebido enriquecido del método propuesto logra una separación más definida, donde los nódulos benignos y malignos tienden a agruparse con mínima superposición, lo que sugiere una mayor capacidad discriminativa para la clasificación.



**Figura 5.** Comparación de la distribución del embebido generado por la U-Net con atención (1) y por el método propuesto (2).

## 6. CONCLUSIONES Y TRABAJO FUTURO

En este trabajo, se presentó una metodología para la clasificación de nódulos pulmonares basada en un enfoque semi-supervisado que integra información geométrica, morfológica, representaciones profundas extraídas mediante una arquitectura basada en la U-Net con mecanismos de atención cruzada. Se evaluó la capacidad discriminativa de los descriptores geométricos y de las anotaciones morfológicas, demostrando que contienen información relevante para la clasificación de nódulos pulmonares. Además, se realizó un experimento exhaustivo de diferentes configuraciones del modelo, desde el uso de la U-Net sin atención hasta la integración completa de todas las representaciones. Como resultado, se obtuvieron métricas de desempeño competitivas, alcanzando un AUC de 93.02%, una especificidad de 91.18% y un F1-score de 95.32%. Estos resultados demuestran la eficacia del método propuesto y su aporte para mejorar la clasificación de nódulos pulmonares.

Para llevar a cabo el estudio se utilizó el conjunto de datos público LIDC-IDRI. En particular, se consideraron 530 nódulos con etiquetas de benigno 174 o maligno 356, excluyendo aquellos clasificados como indeterminados. Se diseñó entonces una metodología que permite capturar la forma del nódulo en términos de geometría y morfología con el fin de obtener un embebido profundo robusto que permita obtener una clasificación confiable. La comparación con trabajos representativos del estado del arte confirma la competitividad del enfoque propuesto. En particular, los enfoques basados en características radiómicas y descriptores morfológicos han cobrado especial relevancia por su capacidad de capturar propiedades que los radiólogos ya reconocen como indicios de malignidad, tales como la espiculación o la lobulación. Por ejemplo, el trabajo de Tang

*et al.*<sup>47</sup> resalta el valor de los descriptores cuantitativos clásicos al combinarlos con estrategias de ensemble learning, lo cual permite una representación diversa de los nódulos a partir de medidas radiómicas de forma e intensidad. Ma *et al.*<sup>48</sup> llevan esta idea aún más lejos al integrar múltiples redes convolucionales profundas y modelos de grafos junto con atributos radiómicos, generando una arquitectura de fusión compleja que captura relaciones estructurales avanzadas entre las características. Por otro lado, en términos semi-supervisados, el trabajo de Apostolopoulos *et al.*<sup>49</sup> propone una estrategia semi-supervisada que combina redes generativas adversarias (GANs) para generar ejemplos sintéticos realistas, con una arquitectura VGG19 para extraer características discriminativas desde imágenes con etiquetas limitadas. Adicionalmente, Shi *et al.*<sup>50</sup> presenta un enfoque también semi-supervisado, donde se utiliza transferencia de aprendizaje y una estrategia iterativa de selección de muestras no etiquetadas para refinar el clasificador, basándose en la similitud de las representaciones latentes. Teniendo en cuenta esto, los resultados obtenidos muestran diferencias relevantes. Tang *et al.* reportaron un AUC de 89.53%, lo cual refleja la utilidad de los atributos radiómicos, aunque al compararlo con el método propuesto, que obtuvo un AUC de 93.02%, se evidencia una mejora significativa. Esto sugiere que la integración con descriptores geométricos y profundos permite una caracterización más completa del nódulo. De igual manera, *Apostolopoulos et al.*

---

<sup>47</sup> Guozhi TANG et al. “Multi-type classification of lung nodules based on CT radiomics and ensemble learning for diversity weighting”. In: *Quantitative Imaging in Medicine and Surgery* 14.12 (2024), p. 8942.

<sup>48</sup> Ling MA et al. “A novel fusion algorithm for benign-malignant lung nodule classification on CT images”. In: *BMC Pulmonary Medicine* 23.1 (2023), p. 474.

<sup>49</sup> Ioannis D APOSTOLOPOULOS; Nikolaos D PAPATHANASIOU, and George S PANAYIOTAKIS. “Classification of lung nodule malignancy in computed tomography imaging utilising generative adversarial networks and semi-supervised transfer learning”. In: *Biocybernetics and Biomedical Engineering* 41.4 (2021), pp. 1243–1257.

<sup>50</sup> Feng SHI et al. “Semi-supervised deep transfer learning for benign-malignant diagnosis of pulmonary nodules in chest CT images”. In: *IEEE transactions on medical imaging* 41.4 (2021), pp. 771–781.

alcanzaron un AUC de 92.08% y un F1-score de 91.9%, indicando una alta sensibilidad del modelo al incorporar muestras generadas artificialmente. Sin embargo, nuestro método lo supera tanto en AUC (93.02%) como en F1-score (95.32%), lo cual sugiere que la incorporación de información estructural real —tanto geométrica como morfológica— ofrece una ventaja frente a estrategias que dependen de datos sintéticos. En el caso de Shi *et al.*, el modelo alcanzó un AUC de 91.0% y un F1-score de 91.9%, evidenciando el aporte del aprendizaje semi-supervisado cuando se dispone de pocos datos etiquetados. No obstante, el enfoque propuesto demuestra una mayor consistencia en sus métricas, superando estos valores y mostrando un equilibrio superior entre sensibilidad y precisión. Ma *et al.* lograron un AUC de 96.29% y un F1-score de 91.14% al combinar radiomics con deep learning, evidenciando un rendimiento alto en la clasificación. Incluso y aún más importante, el método propuesto evita sesgos en la malignidad siendo competitiva en estado del arte.

Los resultados obtenidos en este trabajo evidencian el potencial de integrar descriptores profundos de imagen con información radiológica y geométrica para la clasificación robusta de nódulos pulmonares. En particular, se mostró que la combinación de características de distinto origen proporciona una representación discriminativa rica, capaz de superar enfoques tradicionales que se basan únicamente en características Radiomics o en modelos de deep learning puros. Sin embargo, este estudio presenta algunas limitaciones importantes, entre ellas el tamaño reducido del conjunto de datos; adicionalmente, el uso de cortes 2D axiales para la segmentación y geometría (en lugar de volúmenes 3D completos) podría estar simplificando en exceso la anatomía del nódulo y obviando información contextual del pulmón. Como trabajo futuro, se propone explorar nuevos conjuntos de datos en los que se incluya información de biopsia u otros datos clínicos; esto permitiría validar la robustez del modelo con diagnósticos confirmados y añadir contexto clínico al análisis. Por último, la incorporación de nuevas fuentes de información tanto geométri-

cas como morfológicas; por ejemplo, descriptores de forma más complejos, parámetros radiológicos adicionales u otros biomarcadores extraídos de las imágenes representa una vía prometedora para potenciar aún más el desempeño del sistema. Adicionalmente, la incorporación de enfoques de aprendizaje auto-supervisado o el pre-entrenamiento en grandes bases de datos no anotadas podría generar representaciones profundas más robustas sin depender tanto de etiquetas manuales.

## BIBLIOGRAFÍA

ABERLE, Denise R et al. “Results of the two incidence screenings in the National Lung Screening Trial”. In: *New England Journal of Medicine* 369.10 (2013), pp. 920–931 (cit. on p. 14).

APOSTOLOPOULOS, Ioannis D; PAPATHANASIOU, Nikolaos D, and PANAYIOTAKIS, George S. “Classification of lung nodule malignancy in computed tomography imaging utilising generative adversarial networks and semi-supervised transfer learning”. In: *Bio-cybernetics and Biomedical Engineering* 41.4 (2021), pp. 1243–1257 (cit. on pp. 41, 42, 45).

ARMATO III, Samuel G et al. “The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans”. In: *Medical physics* 38.2 (2011), pp. 915–931 (cit. on p. 32).

ASTARAKI, Mehdi et al. “Benign-malignant pulmonary nodule classification in low-dose CT with convolutional features”. In: *Physica Medica* 83 (2021), pp. 146–153 (cit. on p. 12).

CRUICKSHANK, Ashleigh; STIELER, Geoff, and AMEER, Faisal. “Evaluation of the solitary pulmonary nodule”. In: *Internal Medicine Journal* 49.3 (2019), pp. 306–315 (cit. on pp. 14, 16).

CUENCA, Jorge Juan Suárez. “Desarrollo de un sistema de diagnóstico asistido por computador para detección de nódulos pulmonares en tomografía computarizada multicorte.” In: (2009) (cit. on p. 10).

DA NOBREGA, Raul Victor M et al. “RETRACTED ARTICLE: Lung nodule malignancy classification in chest computed tomography images using transfer learning and convolutional neural networks”. In: *Neural Computing and Applications* 32.15 (2020), pp. 11065–11082 (cit. on p. 33).

DEY, Raunak; LU, Zhongjie, and HONG, Yi. “Diagnostic classification of lung nodules using 3D neural networks”. In: *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*. IEEE. 2018, pp. 774–778 (cit. on p. 33).

“Estudio y Seguimiento del Nódulo Pulmonar Solitario Detectado por RX y TC”. In: (2023) (cit. on p. 10).

FU, Xiaohang et al. “Attention-enhanced cross-task network for analysing multiple attributes of lung nodules in ct”. In: *arXiv preprint arXiv:2103.03931* (2021) (cit. on pp. 11, 21).

HAMIDIAN, Sardar et al. “3D convolutional neural network for automatic detection of lung nodules in chest CT”. In: *Medical Imaging 2017: Computer-Aided Diagnosis*. Vol. 10134. SPIE. 2017, pp. 54–59 (cit. on p. 21).

HUA, Kai-Lung et al. “Computer-aided classification of lung nodules on computed tomography images via deep learning technique”. In: *OncoTargets and therapy* (2015), pp. 2015–2022 (cit. on p. 10).

HUSSEIN, Sarfaraz et al. “Tumornet: Lung nodule characterization using multi-view convolutional neural network with gaussian process”. In: *2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017)*. IEEE. 2017, pp. 1007–1010 (cit. on p. 11).

INTERNATIONAL AGENCY FOR RESEARCH ON CANCER. *Global Cancer Observatory*. <https://gco.iarc.fr/>. 2022 (cit. on pp. 10, 14, 23).

KASTNER, Julia et al. “Lung-RADS version 1.0 versus lung-RADS version 1.1: comparison of categories using nodules from the national lung screening trial”. In: *Radiology* 300.1 (2021), pp. 199–206 (cit. on p. 15).

KHAN, Muhammad Attique et al. “VGG19 network assisted joint segmentation and classification of lung nodules in CT images”. In: *Diagnostics* 11.12 (2021), p. 2208 (cit. on p. 22).

LI, Rui et al. “Deep learning applications in computed tomography images for pulmonary nodule detection and diagnosis: A review”. In: *Diagnostics* 12.2 (2022), p. 298 (cit. on p. 19).

MA, Ling et al. “A novel fusion algorithm for benign-malignant lung nodule classification on CT images”. In: *BMC Pulmonary Medicine* 23.1 (2023), p. 474 (cit. on pp. 41, 42, 45).

MAHMOOD, Sozan Abdullah and AHMED, Hunar Abubakir. “An improved CNN-based architecture for automatic lung nodule classification”. In: *Medical & Biological Engineering & Computing* 60.7 (2022), pp. 1977–1986 (cit. on p. 11).

MEJIA, Cecilia Irene Loeza; BISWAL, Rajesh Roshan, and SÁNCHEZ-DELACRUZ, Eddy. “Estado del arte de clasificación de nódulos pulmonares utilizando Aprendizaje Automático”. In: (2019) (cit. on p. 14).

RIEL, Sarah J van et al. “Observer variability for classification of pulmonary nodules on low-dose CT images and its effect on nodule management”. In: *Radiology* 277.3 (2015), pp. 863–871 (cit. on p. 11).

RIQUELME, Diego and AKHLOUFI, Moulay A. “Deep learning for lung cancer nodules detection and classification in CT scans”. In: *Ai* 1.1 (2020), pp. 28–67 (cit. on p. 19).

RONNEBERGER, Olaf; FISCHER, Philipp, and BROX, Thomas. “U-net: Convolutional networks for biomedical image segmentation”. In: *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III* 18. Springer. 2015, pp. 234–241 (cit. on pp. 17, 18).

SALIHÖĞLU, Yavuz Sami et al. “Diagnostic performance of machine learning models based on 18F-FDG PET/CT radiomic features in the classification of solitary pulmonary nodules”. In: *Molecular Imaging and Radionuclide Therapy* 31.2 (2022), p. 82 (cit. on p. 22).

SHAH, Asghar Ali et al. “Deep learning ensemble 2D CNN approach towards the detection of lung cancer”. In: *Scientific Reports* 13.1 (2023), p. 2987 (cit. on p. 20).

SHELHAMER, Evan; LONG, Jonathan, and DARRELL, Trevor. “Fully convolutional networks for semantic segmentation”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.4 (2017), pp. 640–651 (cit. on pp. 18, 19).

SHI, Feng et al. “Semi-supervised deep transfer learning for benign-malignant diagnosis of pulmonary nodules in chest CT images”. In: *IEEE transactions on medical imaging* 41.4 (2021), pp. 771–781 (cit. on pp. 41, 42, 45).

SONG, Jingqi et al. “Weakly-supervised classification of pulmonary nodules based on shape characters”. In: *2016 IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*. IEEE. 2016, pp. 228–232 (cit. on p. 22).

TANG, Guozhi et al. “Multi-type classification of lung nodules based on CT radiomics and ensemble learning for diversity weighting”. In: *Quantitative Imaging in Medicine and Surgery* 14.12 (2024), p. 8942 (cit. on pp. 41, 42, 45).

TANNER, Nichole T et al. “Management of pulmonary nodules by community pulmonologists: a multicenter observational study”. In: *Chest* 148.6 (2015), pp. 1405–1414 (cit. on p. 16).

TONG, Chao et al. “Pulmonary nodule classification based on heterogeneous features learning”. In: *IEEE Journal on Selected Areas in Communications* 39.2 (2020), pp. 574–581 (cit. on p. 20).

TONG, Guofeng et al. “Improved U-NET network for pulmonary nodules segmentation”. In: *Optik* 174 (2018), pp. 460–469 (cit. on p. 17).

URREHMAN, Zia et al. “Effective lung nodule detection using deep CNN with dual attention mechanisms”. In: *Scientific Reports* 14.1 (2024), p. 3934 (cit. on p. 21).

VIDAL, Tamara Laxe et al. “Enfrentándonos a la variabilidad: recomendaciones de la Sociedad Fleischner para medir correctamente los nódulos pulmonares”. In: *Seram* 1.1 (2021) (cit. on pp. 16, 17).

XU, Qing-qing et al. “Prediction efficacy of feature classification of solitary pulmonary nodules based on CT radiomics”. In: *European Journal of Radiology* 139 (2021), p. 109667 (cit. on pp. 12, 22).

YANG, Li; ZHANG, Q., and AL., et. “Assessment of the cancer risk factors of solitary pulmonary nodules”. In: *Oncotarget* 8.17 (2017), p. 29318 (cit. on p. 15).

YU, Hui et al. “Design of lung nodules segmentation and recognition algorithm based on deep learning”. In: *BMC bioinformatics* 22 (2021), pp. 1–21 (cit. on p. 17).

ZHAI, Penghua et al. “Multi-task learning for lung nodule classification on chest CT”. In: *IEEE access* 8 (2020), pp. 180317–180327 (cit. on p. 20).

ZHAO, Xinzhuo et al. “Agile convolutional neural network for pulmonary nodule classification using CT images”. In: *International journal of computer assisted radiology and surgery* 13 (2018), pp. 585–595 (cit. on p. 11).