

CLASIFICACIÓN DE LA MALIGNIDAD DE LOS NÓDULOS PULMONARES DESDE
REPRESENTACIONES GEOMÉTRICAS PROFUNDAS

SANTIAGO LEAL CASANOVA

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

2025

CLASIFICACIÓN DE LA MALIGNIDAD DE LOS NÓDULOS PULMONARES DESDE
REPRESENTACIONES GEOMÉTRICAS PROFUNDAS

SANTIAGO LEAL CASANOVA

Tesis presentada en cumplimiento de los requisitos para optar por el título de:
Ingeniero de Sistemas

Director:

Fabio Martínez Carrillo, Ph.D.

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

2025

DEDICATORIA

A Dios, a la virgencita y a mis padres Adriana Casanova Pabón, Oriol Leal Lemus, a mi nonita Leonor Pabón de Casanova y a mi nana Marina Almeida Esparza. Por ser aquellos que siempre han estado ahí conmigo. No solo brindandome apoyo, sino amor incondicional, guía, y fé, fé en mis capacidades, aptitudes y esa ciega confianza que me depositan todos los días sobre lo alto y lo lejos que puedo llegar en la vida. No podría estar más agradecido con la vida y con Dios por darme la fortuna y la bendición de ser su hijo/nieto/ahijado. Y lo único y mejor que puedo es agradecer, y prometerles que toda esa fé y confianza que me han dado no será en vano. Muchas gracias

Santiago Leal Casanova

AGRADECIMIENTOS

A mi director el profesor Fabio Martinez Carillo, y a mi codirectora Alejandra Moreno Tarazona. Soy consciente que durante mi paso a través del grupo BIVL²ab no he sido el mejor estudiante, tal vez a veces ni siquiera bueno, pero a pesar todo eso y mis pobres avances en la investigación, siempre me recibieron con una sonrisa, y mucha paciencia, y aun mas important, me acogieron como uno de los suyos y me dieron un lugar al cual pertenecer, a la pequeña gran familia BIVL²ab. Aleja fue la persona que guió por el camino de la ciencia, me enseñó la mayoría de cuanto se acerca de todo el conocimiento que hizo posible llevar a cabo esta tesis, y aún con mis limitados avances semana a semana nunca se rindio o renego de mi, y por el contrario siempre me seguia ayudando y recibíendome alegremente aun si no lo merecia; la vida definitivamente me bendijo, cuando de todas las personas posibles, la persona que me toco como codirector fue ella. También quiero destacar y reiterar el apoyo y mis agradecimiento al profe Fabio. No tanto porque sea mi director, quiero decir, no por ser el director significa que tus dirigidos van a apreciarte, pero yo si aprecio mucho al profe Fabio. Él me brindó un apoyo invaluable para poder llegar aquí y entregar esta tesis. Él en ningún momento tenía el porqué ayudarme, de hecho lo más fácil hubiera sido no hacerlo, pero aún así lo hizo y para mí significa mucho. Soy consciente de que sin él, sin su ayuda e intersección, la tesis aquí presente no habría sido posible ni habría podido ser entregada. Profe, si alguna vez llega a leer este documento, muchisimas gracias por todo; le doy mis más sinceros agradecimientos.

A mis compañeros del grupo de investigación BIVL²ab, gracias a cada uno de ustedes la ciencia fue mucho más divertida con su compañía y fueron ustedes y esta familia la que hizo que me enamorara por la ciencia.

A Dios, a la virgen y a mi familia, por ser el pilar que siempre estuvo sosteniéndome en todo momento. Soy quien soy gracias a ustedes y en esta vida podré no saber muchas cosas, pero si algo sé es que no podría ser más afortunado de tenerlos en mi vida.

A mis amigos, en especial a Laurita, Oscar, Juan David y Daniel, fueron un refugio en estos últimos cinco años de universidad. Gracias por estar a mi lado, brindarme su apoyo, y compartir conmigo todas esas pequeñas victorias hasta llegar aquí, hasta el final.

Santiago Leal Casanova

CONTENIDO

	pág.
INTRODUCCIÓN	12
Justificación de cursos de programas de maestría	16
0.1. Matemáticas avanzadas	16
0.2. Seminario de Investigación	17
1. FUNDAMENTOS Y TRABAJOS PREVIOS	18
1.1. Cáncer de pulmón	18
1.2. Metodologías computacionales	19
1.2.1 Métodos basados en convoluciones	19
1.2.2 Metodos basados en atención	21
1.2.3 Métodos multiescala	22
1.3. Caracterización de nódulos pulmonares desde enfoques locales y contextuales . .	23
2. PROBLEMA DE INVESTIGACIÓN	26
3. OBJETIVOS	27
3.1. Objetivo general	27
3.2. Objetivos Específicos	27
4. MÉTODO PROPUESTO	28
4.1. Codificación de atención del nódulo	29
4.2. Una representación contextual 3D a escala pulmonar	30
4.3. Diseño experimental	32
5. EVALUACIÓN Y RESULTADOS	34

6. CONCLUSIONES Y TRABAJO FUTURO	39
BIBLIOGRAFÍA	42

LISTA DE FIGURAS

	pág.
Figura 1. Arquitectura convolucional	20
Figura 2. Arquitectura de atención	21
Figura 3. Pipeline del metodo propuesto	29
Figura 4. Gráfica del impacto del bloque tranformer	35
Figura 5. Figura de embebidos obtenidos	37
Figura 6. Figura de embebidos obtenidos	38

LISTA DE TABLAS

	pág.
Tabla 1. Estudio de ablación del enfoque propuesto	34
Tabla 2. Comparación de resultados frente al estado del arte.	36

RESUMEN

TÍTULO: Clasificación de la malignidad de los nódulos pulmonares desde representaciones geométricas profundas *

AUTOR: Santiago Leal Casanova. **

PALABRAS CLAVE: Cáncer de pulmón, nódulos pulmonares, información contextual, aprendizaje multiescala.

DESCRIPCIÓN: El cáncer de pulmón (CP) es el cáncer con mayor mortalidad. Los nódulos pulmonares (NP) son los principales biomarcadores del cáncer pulmonar, siendo masas anormales de entre 3 y 30 mm, categorizadas según observaciones radiológicas de expertos. Sin embargo, este análisis es subjetivo, por lo que existen métodos computacionales para brindar un soporte a la clasificación de los NP. Estos métodos principalmente se centran en la codificación local del nódulo, acotada a la región de observación, perdiendo características contextuales, relevantes en la clínica para dar un mayor soporte. En este contexto de investigación, durante la modalidad de “Cursos en Programas de Maestría”. Teniendo en cuenta el aporte de las asignaturas y la consolidación de la investigación, en este trabajo se desarrollo un modelo basado en transformers que integra observaciones a escala local de nódulos 3D y contexto a escala pulmonar para mejorar la clasificación de malignidad. El contexto pulmonar se codifica a partir de un codificador de transformer estándar, proporcionando un vector de con información contextual robusta. Paralelamente, un codificador de atención multicabeza especializado captura la información local a escala del nódulo para mejorar la clasificación del nódulo, que posteriormente se integra con la información contextual. El método propuesto se validó en el conjunto de datos LIDC-IDRI, logrando un 96,44% y un 95,53% de AUC y recall respectivamente, superando a los métodos más avanzados.

* Trabajo de grado

** Facultad de Ingenierías Físicomecánicas. Escuela de Ingeniería de Sistemas e Informática. Director: Fabio Martínez, PhD.

ABSTRACT

TITLE: Classification of malignancy in pulmonary nodules using deep geometric representations *

AUTHOR: Santiago Leal Casanova. **

KEYWORDS: Lung cancer, lung nodules, contextual information, multiscale learning.

DESCRIPTION: Lung cancer (PC) is the cancer with the highest mortality. Lung nodules (PN) are the main biomarkers of lung cancer, being abnormal masses between 3 and 30 mm, categorized according to expert radiological observations. However, this analysis is subjective, so computational methods exist to support the classification of PNs. These methods mainly focus on the local coding of the nodule, limited to the region of observation, losing contextual features, relevant in the clinic to provide further support. In this research context, during the modality of “Courses in Master Programs”. Taking into account the contribution of the courses and the consolidation of the research, this work developed a transformer-based model that integrates local-scale observations of 3D nodules and lung-scale context to improve the classification of malignancy. The lung context is encoded from a standard transformer encoder, providing a vector of robust contextual information. In parallel, a specialized multi-head attention encoder captures local nodule-scale information to improve nodule classification, which is then integrated with the contextual information. The proposed method was validated on the LIDC-IDRI dataset, with 96.44% and 95.53% AUC and recall respectively, surpassing the most advanced methods.

* Undergraduate thesis

** Faculty of Physics-Mechanics Engineering. School of Systems Engineering and Informatics. Advisor: Fabio Martínez, PhD.

INTRODUCCIÓN

El cáncer de pulmón (CP) es la principal causa de mortalidad por cáncer a nivel mundial. Únicamente en el año 2022, se registró una alarmante cifra de aproximadamente 2,4 millones de nuevos casos y 1,8 millones de muertes causadas por esta enfermedad a nivel mundial. En Colombia, las cifras continúan siendo alarmantes: el cáncer de pulmón es el quinto con mayor incidencia y representa el 10,8% de las muertes causadas por cáncer ¹. La tomografía computarizada (TC) es el principal medio de apoyo para el análisis del CP, ya que proporciona una visualización completa de la morfología pulmonar. Esto permite a los radiólogos identificar y caracterizar anomalías que pueden ser indicativas de cáncer pulmonar, así como realizar el seguimiento del paciente ^{2 3}.

Los nódulos pulmonares son los principales biomarcadores del cáncer de pulmón. La Sociedad de Fleischner los define como opacidades redondeadas, más o menos bien definidas, que miden entre 3 y 30 mm. ⁴ Mediante las secuencias de TC, es posible caracterizar los nódulos pulmonares analizando características como el tamaño y la morfología del nódulo, que incluye rasgos tales como la lobulación, espiculación, densidad y textura, entre otros. Además, se pueden observar características contextuales, como las estructuras adyacentes al

¹ International Agency for RESEARCH ON CANCER (IARC). *Global Cancer Observatory*. Recuperado de IARC. 2022.

² American Cancer SOCIETY. *Detección del cáncer de pulmón*. Accedido el 25 de noviembre de 2024. 2023.

³ National Lung Screening Trial Research TEAM. “Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening”. In: *New England Journal of Medicine* 378.24 (2017), pp. 2285–2296. DOI: [10.1056/NEJMoa1613682](https://doi.org/10.1056/NEJMoa1613682).

⁴ Konstantinos LOVERDOS et al. “Lung nodules: a comprehensive review on current approach and management”. In: *Annals of Thoracic Medicine* 14.4 (2019), pp. 226–238.

nódulo, su posición y las diferentes lesiones o estructuras presentes en el pulmón. La caracterización de un nódulo pulmonar suele realizarse de manera observacional por radiólogos expertos siguiendo estándares como el Lung-RADS. Sin embargo, este esquema es subjetivo, lo que provoca desacuerdos tanto a nivel inter-observador como intra-observador^{5 6 7}.

Con el objetivo de realizar una herramienta de soporte a la clasificación de los nódulos pulmonares, se han desarrollado múltiples métodos para caracterizar la malignidad de los mismos. Por un lado, existen los métodos basados en redes convolucionales, que a partir de la imagen y/o volumen del nódulo generan mapas de características que contienen información profunda del mismo^{8 9 10 11}. Por otro lado, están los métodos basados en atención,

-
- ⁵ S. 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.
- ⁶ Rabia JAVED et al. “Deep learning for lungs cancer detection: a review”. In: *Artificial Intelligence Review* 57.8 (2024), p. 197. DOI: [10.1007/s10462-023-10234-5](https://doi.org/10.1007/s10462-023-10234-5).
- ⁷ Alina Cornelia PACURARI et al. “Diagnostic accuracy of machine learning AI architectures in detection and classification of lung cancer: a systematic review”. In: *Diagnostics* 13.13 (2023), p. 2145.
- ⁸ Mundher AL-SHABII et al. “Gated-Dilated Networks for Lung Nodule Classification in CT scans”. In: *IEEE Access* 7 (2019), pp. 178827–178838.
- ⁹ Fuyan LIAO et al. “Efficient Lung Nodule Classification Using Transferable Texture Convolutional Neural Network”. In: *IEEE Access* 8 (2020), pp. 175859–175870.
- ¹⁰ et. al YI WANG. “Novel Convolutional Neural Network Architecture for Improved Pulmonary Nodule Classification on Computed Tomography”. In: *Multidimensional Systems and Signal Processing* 31 (2020), pp. 1163–1183.
- ¹¹ Muhammad MUZAMMIL et al. “Pulmonary Nodule Classification Using Feature and Ensemble Learning-Based Fusion Techniques”. In: *IEEE Access* 9 (2021), pp. 113415–113427.

capaces de capturar las relaciones de largo alcance entre las diferentes partes del nódulo^{12 13}
^{14 15}. Sin embargo, estos métodos suelen analizar únicamente la imagen del nódulo pulmonar, lo que limita la información que el modelo puede obtener. Además, suelen ignorar características contextuales, como su posición o las estructuras circundantes, que médicamente han demostrado ser significativas para la clasificación de su malignidad.

Bajo el presente contexto, y en coherencia con la modalidad de “Cursos en Programas de Maestría”, se realizaron cursos que aportaron conocimientos de relevancia para realizar una investigación en el marco de los modelos de inteligencia artificial para el análisis del cáncer de pulmón, que permitan integrar información local de nódulo con información contextual del pulmón para la estratificación de malignidad de nódulos pulmonares. Para lograr dicho objetivo y bajo la dirección del director de tesis, se dictaminó que los cursos más adecuados para con la presente propuesta de investigación son los cursos de Matemáticas Avanzadas y Seminario de Investigación. La materia de Matemáticas Avanzadas se escogió en reconocimiento a la importancia de una sólida fundamentación matemática, que sea capaz de dotar de la capacidad de entender los conceptos de aprendizaje profundo, modelos multiescala y de computación necesarios para abordar la investigación propuesta. La materia de Seminario de investigación se eligió con el propósito de instruirse acerca del debido proceso para llevar a

¹² Mundher AL-SHABI et al. “3D axial-attention for lung nodule classification”. In: *International Journal of Computer Assisted Radiology and Surgery* 16 (2021), pp. 1319–1324.

¹³ Alejandra MORENO et al. “A volumetric multi-head attention strategy for lung nodule classification in CT”. in: *Medical Imaging 2023: Computer-Aided Diagnosis*. Vol. 12465. SPIE, 2023, p. 1246504. DOI: [10.1117/12.2653270](https://doi.org/10.1117/12.2653270).

¹⁴ et.al KAI XIA. “Adaptive Aggregated Attention Network for Pulmonary Nodule Classification”. In: *Applied Sciences* 11.2 (2021), p. 610.

¹⁵ Hossein Karshenas AHMED SAIHOOD and Ahmad Reza NAGHSH-NILCHI. “Multi-Orientation Local Texture Features for Guided Attention-Based Fusion in Lung Nodule Classification”. In: *IEEE Access* 11 (2023), pp. 17555–17568.

cabo una investigación.

En esta línea de investigación se desarrolló un novedoso modelo basado en transformers 3D que integra información contextual global con análisis de características localizadas para clasificar de forma robusta la malignidad de los nódulos pulmonares. En general, el enfoque aprovecha los datos volumétricos 3D para extraer características complementarias a escala pulmonar (contextuales) y a escala de nódulo (locales), que posteriormente se fusionan para formar una incrustación integral para la discriminación de malignidad. En concreto, la arquitectura consta de dos ramas: una rama contextual que emplea un transformador estándar para captar las claves posicionales y contextuales de todo el pulmón, y una rama de nódulo que utiliza un bloque de atención espacial especializado para explotar los patrones locales y regionales alrededor del nódulo. La fusión de estos dos descriptores da como resultado una incrustación que distingue eficazmente los nódulos malignos de los benignos. Un enfoque integrado que combina el análisis morfológico y contextual permite una evaluación más precisa del riesgo de malignidad y una mejor planificación del tratamiento.

Justificación de cursos de programas de maestría

En el contexto del presente libro, se fundamenta en la necesidad imperante de adquirir una base sólida de conocimientos avanzados en computación e investigación, lo cual es esencial para garantizar un desarrollo adecuado y de alto nivel de calidad de la propuesta. La complejidad y los desafíos inherentes al estudio de la caracterización de nódulos pulmonares mediante técnicas de aprendizaje profundo requieren una formación especializada que abarque tanto aspectos teóricos como prácticos de la computación moderna y las metodologías de investigación científica.

Con el objetivo de obtener los conocimientos mencionados anteriormente, se ha optado por la modalidad de grado de “Cursos en Programas de Maestría”. Esta decisión está motivada por las ventajas inherentes que ofrece esta modalidad, la cual permite ahondar en áreas especializadas de alto nivel en computación antes de iniciar la fase de investigación propiamente dicha. Los cursos seleccionados están diseñados para proporcionar una comprensión profunda de temas clave sobre las bases teóricas necesarias para comprender las metodologías y arquitecturas más recientes del estado del arte en computación y aprendizaje profundo, así como la estructuración y formalización adecuada de una investigación. Por esto mismo, la modalidad de cursos en programas de maestría dota al estudiante de herramientas y competencias necesarias para abordar eficazmente los desafíos presentes en una propuesta de investigación.

0.1. Matemáticas avanzadas

La asignatura de “Matemáticas Avanzadas” tiene como objetivo principal proporcionar las bases matemáticas necesarias para comprender los conceptos fundamentales del aprendizaje profundo, los modelos multiescala y las técnicas de computación avanzadas. Estos conocimientos son esenciales para abordar de manera efectiva la presente temática de investigación,

ya que permiten al estudiante entender y aplicar herramientas matemáticas clave que son pilares fundamentales para posibilitar entendimiento de los algoritmos y arquitecturas de aprendizaje profundo pertenecientes al estado del arte. Además, el dominio de estos conceptos matemáticos no solo facilita la comprensión teórica de los modelos, sino que también permite su implementación práctica y el análisis de los resultados obtenidos. Esta formación es indispensable para enfrentar los desafíos inherentes al estudio de problemas complejos, como la caracterización de nódulos pulmonares mediante técnicas avanzadas de aprendizaje profundo, asegurando un enfoque riguroso y científicamente sólido en la investigación.

0.2. Seminario de Investigación

Cursar la materia de “seminario de investigación I” tiene el propósito de iniciar al estudiante dentro del mundo de la escritura científica. Esto mediante la enseñanza sobre tópicos necesarios para llevar a cabo el debido proceso a realizar para elaborar adecuadamente una investigación. Se espera que al tomar el curso seminario de investigación I este aporte conocimientos esenciales acerca de la correcta formulación de una investigación, así como el aprendizaje acerca de apropiada estructura y redacción de un artículo científico. Esta capacitación en formulación y redacción es fundamental para ser dotado con la capacidad construir apropiadamente la formalización y divulgación de los resultados correspondientes a la investigación, las cuales resultan de indiscutible importancia en la apropiada formación de un investigador.

1. FUNDAMENTOS Y TRABAJOS PREVIOS

1.1. Cáncer de pulmón

El cáncer de pulmón es la principal causa de muerte con respecto a los diferentes tipos de cáncer a nivel mundial. Tan solo en el año 2022 se estima que un aproximado de 2.4 millones de personas a lo largo del planeta perdieron la vida a causa de esta enfermedad, además de presentar 2.4 millones de nuevas incidencias de esta enfermedad, representando el 12.4% de todos los nuevos casos de cáncer registrados ¹. Los nódulos pulmonares son masas anormales presentes en el tejido pulmonar con un tamaño que oscila alrededor de los [3-30 milímetros] y son considerados uno de los principales biomarcadores del cáncer pulmonar ⁴. Los nódulos pulmonares son principalmente observados y analizados usando la tomografía computarizada la cual permite observar la morfología completa del pulmón, lo que permite así mismo realizar una caracterización de la morfología del nódulo pulmonar y de su respectivo análisis de malignidad. El análisis de los nódulos pulmonares se basa en aspectos tales como su tamaño, textura, densidad, posición y estructuras circundantes al nódulo, así como las condiciones del tejido pulmonar ^{2 3}.

El análisis de la malignidad de los nódulos pulmonares se basa en una evaluación detallada tanto de sus características morfológicas locales como del contexto pulmonar en el que se encuentran. En el análisis morfológico, se consideran aspectos como el tamaño del nódulo, ya que nódulos mayores de 8 mm presentan un mayor riesgo de ser malignos. La forma y los bordes también son indicadores clave: nódulos con bordes espiculados o irregulares suelen ser más sospechosos que aquellos con bordes lisos y definidos. La densidad del nódulo, que puede ser sólida, semisólida o presentar opacidades en vidrio deslustrado, proporciona información sobre su naturaleza. Además, la textura interna y la presencia de calcificaciones o cavitaciones son factores que ayudan a diferenciar entre lesiones benignas y malignas ^{5 6 7}. El análisis contextual implica examinar la posición del nódulo dentro del pulmón y su relación

con las estructuras circundantes. La ubicación del nódulo, ya sea en las zonas periféricas o centrales del pulmón, puede influir en la interpretación diagnóstica. La proximidad a vasos sanguíneos, bronquios u otras estructuras anatómicas puede aumentar la sospecha de malignidad o afectar las opciones de tratamiento. Asimismo, las condiciones del tejido pulmonar circundante, como la presencia de enfisema, fibrosis u otras lesiones, pueden alterar la apariencia del nódulo y complicar su evaluación. Un enfoque integral que combine el análisis morfológico y contextual permite a los especialistas determinar con mayor precisión el riesgo de malignidad y planificar el manejo clínico adecuado ^{4 16 17}.

1.2. Metodologías computacionales

En el ámbito del análisis médico, el uso de metodologías computacionales para el apoyo del diagnóstico, ha sido clave en aras de mejorar la detección y clasificación de los nódulos pulmonares. En el estado del arte, se han empleado diversas arquitecturas y modelos, que han mejorado la precisión y eficacia de la clasificación de malignidad de los nódulos pulmonares. Estos enfoques capturan características y patrones de relevancia sobre las imágenes médicas. Algunas de las arquitecturas más efectivas hasta la fecha que se escribe este documento usan modelos convolucionales, mecanismos de atención y arquitecturas multiescala, las cuales serán explicadas a continuación.

1.2.1. Métodos basados en convoluciones En los últimos años, los enfoques basados en aprendizaje profundo han surgido como una solución prometedora para mejorar la precisión y consistencia en la clasificación de nódulos pulmonares. Los enfoques convolucionales, en particular las redes neuronales convolucionales (CNNs), han sido fundamentales

¹⁶ Michael K. GOULD et al. “Evaluation of patients with pulmonary nodules: when is it lung cancer?: ACCP evidence-based clinical practice guidelines”. In: *Chest* 132.3 (2007), 108S–130S.

¹⁷ Helen T. WINER-MURAM. “The solitary pulmonary nodule”. In: *Radiology* 239.1 (2006), pp. 34–49.

en el avance del estado del arte en este dominio debido a su capacidad para extraer automáticamente características relevantes de las imágenes médicas. Las CNNs son especialmente efectivas en el procesamiento de imágenes, ya que pueden capturar patrones locales y complejos a través de sus capas convolucionales y de agrupamiento, lo que permite una mejor discriminación entre nódulos benignos y malignos, mejorando la precisión diagnóstica y reduciendo la variabilidad en la interpretación humana.

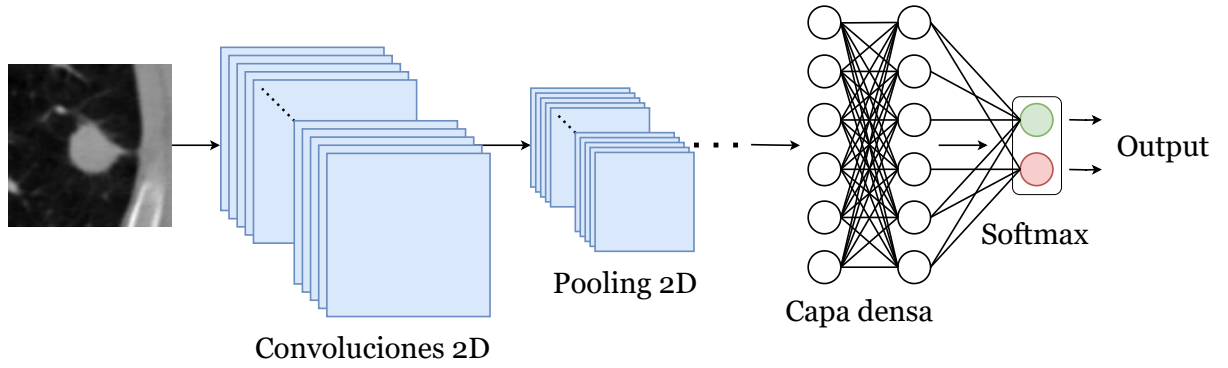


Figura 1. Ilustración de una arquitectura convolucional usada para la clasificación de nódulos pulmonares.

En la caracterización de nódulos pulmonares, las convoluciones juegan un papel fundamental al aplicar filtros o núcleos a un conjunto de imágenes recortadas de los nódulos, denotado como $X = \{x_1, x_2, \dots, x_n\}$, donde cada $x_i \in \mathbb{R}^{H \times W \times C}$ representa una imagen con altura H , anchura W y C canales de color. Cada filtro $f_j \in \mathbb{R}^{k_H \times k_W \times C}$ recorre la imagen de entrada realizando operaciones matemáticas que extraen características específicas, como bordes, texturas y formas. La operación de convolución entre una imagen x_i y un filtro f_j se define como:

$$(x_i * f_j)(m, n) = \sum_{u=0}^{k_H-1} \sum_{v=0}^{k_W-1} \sum_{c=1}^C x_i(m+u, n+v, c) \cdot f_j(u, v, c)$$

Este proceso genera un mapa de características $M_{ij} \in \mathbb{R}^{H' \times W'}$, donde $H' = H - k_H + 1$ y $W' = W - k_W + 1$. Al aplicar múltiples capas convolucionales y de agrupamiento

(pooling), la red neuronal es capaz de aprender representaciones cada vez más profundas y complejas de las características de los nódulos pulmonares. Estas representaciones profundas son esenciales para tareas de clasificación, ya que permiten al modelo capturar patrones intrincados asociados con la malignidad o benignidad de los nódulos.

1.2.2. Metodos basados en atención A pesar del éxito de los métodos convolucionales presentan limitaciones. Una de las principales es su incapacidad para capturar relaciones a larga distancia y dependencias globales en las imágenes, ya que las convoluciones operan principalmente en regiones locales. Esto puede resultar en la pérdida de información contextual importante que podría influir en la precisión de la clasificación. Además, las CNNs pueden ser propensas al sobreajuste, especialmente cuando se entrenan con conjuntos de datos limitados o desequilibrados, como es común en el ámbito médico. Estas limitaciones han motivado la exploración de enfoques alternativos y la incorporación de técnicas adicionales para mejorar el rendimiento y la generalización de los modelos de clasificación de nódulos pulmonares.

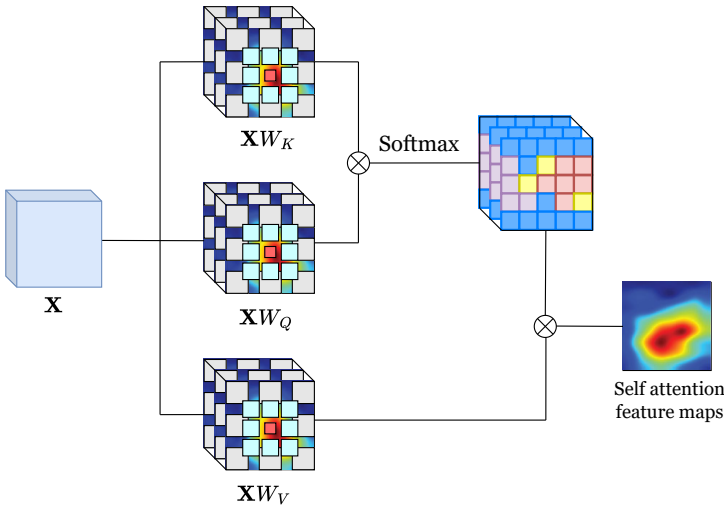


Figura 2. Ilustración de una arquitectura que usa mecanismos de atención usada para la clasificación de nodulos pulmonares.

Los mecanismos de atención han ganado popularidad en el procesamiento de imágenes médicas por su capacidad para modelar relaciones globales y capturar dependencias a larga distancia. Hablando en el dominio de los nódulos pulmonares, los mecanismos de atención permiten relacionar diferentes partes del nódulo que no necesariamente se encuentran adyacentes espacialmente pero que comparten una relación que puede ser indicativa de malignidad. Estos mecanismos de atención funcionan asignando diferentes pesos a partes de la entrada, la cual en nuestro problema corresponderá a imágenes de nódulos pulmonares. De esta manera, permitiendo que el modelo se enfoque en las características más relevantes. Específicamente, utilizan tres representaciones: consulta $Q = \mathbf{XW}_Q$, clave $K = \mathbf{XW}_K$ y valor $V = \mathbf{XW}_V$, generados mediante proyecciones lineales de la matriz de características de entrada \mathbf{X} , donde W_Q, W_K, W_V son matrices de pesos aprendidas.

El mecanismo de atención se calcula como una suma ponderada de los valores, con pesos basados en la similitud entre consultas y claves: $Atencion(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V$. Aquí, QK^T es el producto de consultas y transpuestas de claves, $\sqrt{d_k}$ es un factor de normalización correspondiente a la dimensión de las claves, y la función softmax convierte las similitudes en una distribución de probabilidad que actúa como máscara de atención. Este proceso permite capturar relaciones a larga distancia, ya que cada salida es una combinación ponderada de todas las entradas según su relevancia.

1.2.3. Métodos multiescala Además de los métodos anteriores que se basan en la extracción de características nodulares, se han explorado varios métodos multiescala para proporcionar un análisis más completo al considerar información más amplia, incluyendo el contexto de los nódulos pulmonares. Los modelos multiescala ofrecen la ventaja de capturar tanto las características detalladas del nódulo como la información contextual de sus alrededores. Esto permite al modelo obtener información de las estructuras adyacentes al nódulo, como vasos sanguíneos, bronquios y tejido pulmonar circundante, lo cual es crucial para una

evaluación precisa de la malignidad.

En los modelos multiescala enfocados en los nódulos pulmonares, se procesan de manera paralela múltiples escalas del nódulo de forma independiente. Cada escala se analiza mediante una rama convolucional dedicada, obteniendo así diferentes representaciones de las características del nódulo. Denotemos las representaciones obtenidas de cada escala como e_1, e_2, \dots, e_k . Posteriormente, estas representaciones se integran mediante la concatenación de los embeddings de cada rama, lo que se puede expresar matemáticamente como:

$$e = [e_1 \parallel e_2 \parallel \dots \parallel e_k]$$

donde \parallel representa la operación de concatenación. Esta integración permite combinar información de múltiples escalas, enriqueciendo la representación global del nódulo y mejorando la capacidad del modelo para capturar características de malignidad adicionales que en enfoques meramente locales serían imposibles de obtener, con el propósito de mejorar la clasificación de la malignidad de los nódulos pulmonares de manera más precisa.

Sin embargo, los métodos multiescala también presentan limitaciones. Una de las principales es que, en su mayoría, los métodos multiescala aplicados al análisis nodular se limitan a obtener información contextual únicamente de los alrededores inmediatos del nódulo, sin considerar información más lejana. Esto puede restringir la capacidad del modelo para obtener datos sobre la posición global del nódulo dentro del pulmón y las condiciones generales del tejido pulmonar, factores que pueden ser cruciales para la clasificación de la malignidad.

1.3. Caracterización de nódulos pulmonares desde enfoques locales y contextuales

En la caracterización de nódulos pulmonares desde una perspectiva local, diversos métodos basados en mecanismos convolucionales y de atención han demostrado una eficacia significativa. Numerosos estudios han aplicado redes neuronales convolucionales (CNNs) para la clasificación de nódulos pulmonares. Por ejemplo, Al-Shabi et al. (2019) utilizaron con-

voluciones dilatadas para lograr un rendimiento de vanguardia ⁸, mientras que Liao et al. (2019) se enfocaron en extraer patrones texturales de los nódulos ⁹. Wang et al. (2020) combinaron mapas de características de diferentes niveles y núcleos de distintos tamaños para mejorar la clasificación ¹⁰. Además, Muzammil et al. (2021) emplearon técnicas de fusión de características y ensamblaje con clasificadores como SVM y AdaBoostM2, junto a múltiples arquitecturas de CNN (AlexNet, VGG-16, VGG-19), para optimizar la precisión ¹¹. Más recientemente, Chen et al. (2021) propusieron LDNNET, una red densamente conectada con Dense-Block, normalización por lotes y dropout para abordar el desequilibrio de datos y el sobreajuste ¹⁸.

Paralelamente, varios estudios han incorporado mecanismos de atención para mejorar la caracterización de nódulos pulmonares desde una perspectiva local. Al-Shabi et al. (2021) propusieron un modelo de atención axial 3D para reducir la carga computacional ¹², mientras que Moreno et al. (2023) introdujeron una estrategia de atención volumétrica multi-cabeza para capturar patrones complejos ¹³. Xia et al. (2021) utilizaron una red de atención adaptativa para extraer características espaciales y contextuales ¹⁴, y Saihood et al. (2023) combinaron atención con características de textura local multi-orientación ¹⁵. Liu et al. (2024) integraron el microambiente fibroso en la clasificación de malignidad ¹⁹. Sin embargo, los métodos basados en atención presentan limitaciones, como la necesidad de grandes cantidades de datos y potencia computacional debido a su complejidad, además de poder sufrir sobreajuste con conjuntos de datos pequeños o poco diversos. Asimismo, suelen enfocarse únicamente en el nódulo pulmonar, ignorando el contexto circundante, lo que puede omitir características

¹⁸ et. al YING CHEN. “LDNNET: Towards Robust Classification of Lung Nodule and Cancer Using Lung Dense Neural Network”. In: *IEEE Access* 9 (2021), pp. 50301–50320.

¹⁹ Yucheng LIU et al. “Lung Nodule Malignancy Classification with Associated Pulmonary Fibrosis Using 3D Attention-Gated Convolutional Network with CT Scans”. In: *Journal of Translational Medicine* 22.1 (2024), p. 51.

clave para una caracterización completa del nódulo.

2. PROBLEMA DE INVESTIGACIÓN

Los métodos de aprendizaje profundo han demostrado ser herramientas prometedoras para apoyar el diagnóstico, caracterización y clasificación de los nodulos pulmonares. Sin embargo aún existen desafíos en cuanto a la alta variabilidad de los nodulos observados, dificultando la caracterización de la malignidad. De hecho, la mayoría de estos enfoques suelen incluir únicamente la información del nódulo pulmonar, lo cual deriva en que los modelos nunca logren captar la información que corresponde al contexto del nódulo. Además, los métodos multiescala tradicionales suelen operar sobre perspectivas circundantes al nódulo pulmonar, limitándose únicamente a obtener características de malignidad pertenecientes a los alrededores del nódulo, ignorando así estructuras anatómicas o lesiones que pueden ser de utilidad en la clasificación de la malignidad del nódulo pulmonar. Lo anterior contrasta con evidencia en estudios clínicos y radiológicos, donde la información contextual es de gran relevancia, debido a que información como la posición del nódulo, las estructuras presentes en el pulmón o el estado del tejido pulmonar, son ser indicadores relevantes de malignidad.

Estas limitaciones plantean la necesidad de explorar nuevas estrategias que permitan una integración de la información local y contextual en la clasificación de nódulos pulmonares. Surge así la pregunta de investigación:

¿Cómo establecer una arquitectura de aprendizaje profundo que pueda aprovechar eficazmente tanto la información local como la contextual, para clasificar la malignidad de los nódulos pulmonares?

3. OBJETIVOS

3.1. Objetivo general

Implementar una estrategia multiescala profunda para la clasificación de patrones nodulares según su grado de malignidad observado desde las tomografías computarizadas.

3.2. Objetivos Específicos

- Cursar 2 materias de maestría de Matemáticas avanzadas y Seminario de investigación I que soporten la investigación planteada en esta propuesta.
- Recolectar un conjunto de datos de imágenes diagnósticas TC que incluyan anotaciones realizadas por radiólogos sobre la localización y malignidad de nódulos pulmonares.
- Proponer un modelo de aprendizaje profundo que integre las representaciones contextuales para clasificar la malignidad de los nódulos pulmonares.
- Validar la estrategia propuesta respecto a su capacidad para estimar la malignidad de los nódulos pulmonares.

4. MÉTODO PROPUESTO

El presente método introduce una arquitectura que se encarga de analizar tanto la información a escala de nódulo pulmonar como la información a escala del pulmón. Con la información de las dos escalas antes mencionadas, se busca que el modelo sea capaz de capturar las características dicentes de malignidad, tanto a nivel local o nivel de nódulo pulmonar, como también las posibles relaciones que tienen las diferentes estructuras del pulmón con la clasificación del cáncer, para de esta manera poder mejorar la clasificación. Esta arquitectura se compone de dos bloques principales, el primero es el bloque de codificación de atención que se encarga de codificar las características de malignidad que se encuentran en el nódulo pulmonar. Por otra parte, el transformer es el que se encarga de analizar el pulmón en su totalidad con el propósito de capturar las características de las diferentes estructuras del pulmón y las condiciones del tejido pulmonar que puedan estar relacionadas con la clasificación del cáncer. El flujo general del método propuesto se ilustra en la Figura 3.

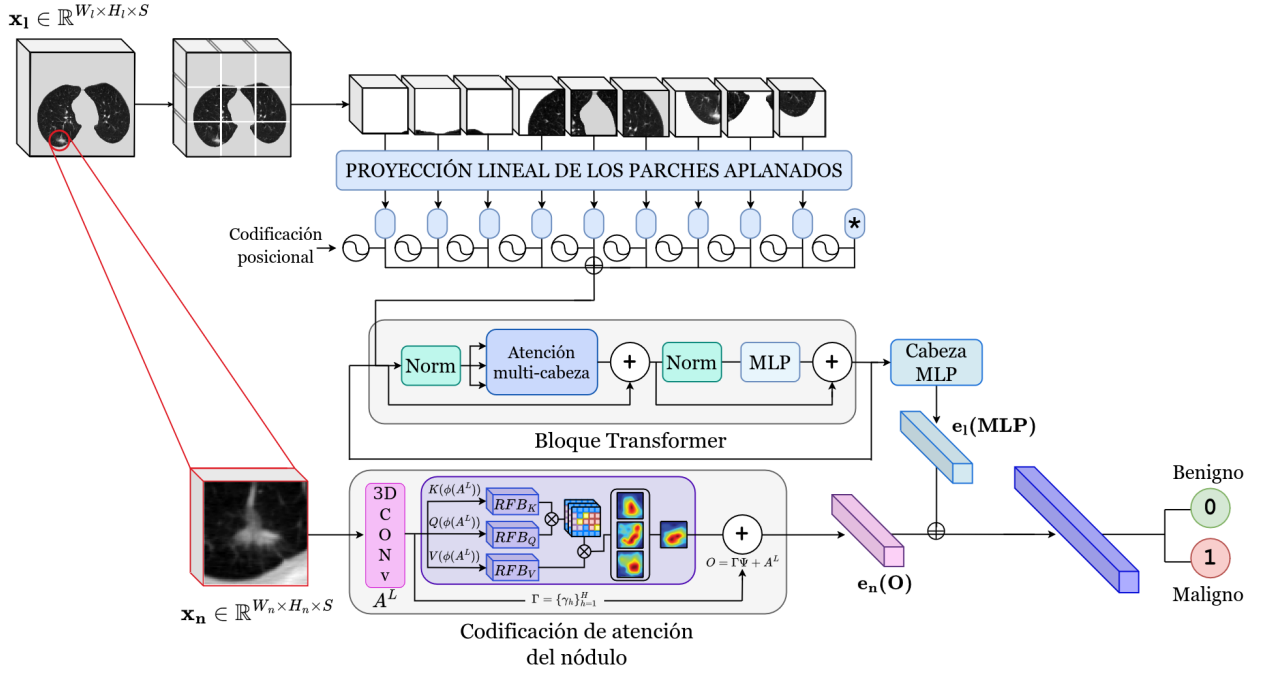


Figura 3. Ilustración de la arquitectura propuesta, donde se integra información a escala del nódulo y a escala del pulmon para clasificar la malignidad del nódulo pulmonar.

4.1. Codificación de atención del nódulo

Aquí, se codifica la información del nódulo a partir de un módulo de atención especializado que comienza capturando características locales de capas convolucionales y luego recupera características no localizadas y regionales a partir de bloques de campo receptivo (RFB). A partir de esta representación enriquecida se extraen características no locales que potencian las relaciones discriminativas, obtenidas mediante un mecanismo de atención espacial. Particularmente, consideramos la codificación volumétrica del nódulo para capturar la compleja morfología, densidad y patrones de textura a lo largo de los ejes. Por lo tanto, la entrada del nódulo es una imagen parche 3D que contiene el nódulo pulmonar (RoI), expresada como un parche $\mathbf{x}_n \in \mathbb{R}^{W_n, H_n, S}$, donde (W_n, H_n) representan las dimensiones espaciales y S el número de cortes.

Luego, cada entrada \mathbf{x}_n se representa como un conjunto de mapas de activación convolu-

cionales $\{A^l\}_{l=1}^L$. A diferencia del método tradicional de autoatención. En el bloque de codificación de atención del nódulo, se consideran proyecciones independientes de RFB debido a la codificación de información intermedia en la imagen para llevar a cabo el mecanismo de autoatención, lo que permite codificar la información intermedia del nódulo en la imagen. Esto se debe a que se utilizan convoluciones dilatadas para cubrir todo el campo receptivo y realzar las características del nódulo. Las proyecciones RFB sobre los mapas convolucionales se representan como $\phi(\mathbb{A}^L)$. Las ramas de autoatención se definen como la clave ($\mathbf{K}(\phi(\mathbf{A}^L)) = \mathbf{W}_k\phi(\mathbf{A}^L)$), la consulta ($\mathbf{Q}(\phi(\mathbf{A}^L)) = \mathbf{W}_q\phi(\mathbf{A}^L)$) y el valor ($\mathbf{V}(\phi(\mathbf{A}^L)) = \mathbf{W}_v\phi(\mathbf{A}^L)$). La matriz de atención se calcula mediante $(m_{ij} = \sigma(\mathbf{K}(\phi(a_i^l))\mathbf{Q}(\phi(a_j^l))))^T$ donde σ denota la función softmax, que asegura que m_{ij} represente pesos de atención normalizados, para establecer la correspondencia no local con el grado de malignidad. Luego, la rama del valor se pondera de acuerdo con la matriz de atención, donde las características de los pesos de atención se calculan mediante $\Psi = \mathbf{M}\mathbf{W}_v\phi(\mathbf{A}^L)$. Aquí, el mecanismo de autoatención RFB se replicó en C cabezas de autoatención en paralelo, que se ajustan recursivamente como $\mathbf{O} = \Gamma\Psi + \mathbf{A}^L$, con un creciente conjunto de parámetros aprendibles $\Gamma = \{\gamma_h\}_{h=1}^H$ que mitigan la contribución del aprendizaje del mecanismo de atención durante las primeras etapas del proceso de entrenamiento, evitando problemas de desvanecimiento del gradiente. Como salida de este modelo de atención, obtenemos un embebido del nódulo $\mathbf{e}_n(\mathbf{O})$ que describe las principales características observadas.

4.2. Una representación contextual 3D a escala pulmonar

Emulando el análisis de un experto, este trabajo también consideró un análisis global a escala pulmonar, aprovechando la categorización contextual según la localización del nódulo y la representación de las estructuras pulmonares circundantes. Para tal caso, la rama a escala pulmonar emplea una metodología basada en ViT que recibe como entrada un volumen pulmonar 3D $\mathbf{x}_1 \in \mathbb{R}^{W_l, H_l, S}$, donde (W_l, H_l) es la dimensión espacial y S indica el número de cortes. Cada tomografía computarizada \mathbf{x}_1 se divide en N parches no solapados, los cuales,

a su vez, se aplanan y se proyectan linealmente en un espacio de embebido de dimensión D , $\{\mathbf{x}_1^n\}_{n=1}^N \in \mathbb{R}^D$. Además, se añade un token de clasificación (*), a los parches de la imagen, el cual, en la salida del último bloque transformer, sirve como representación global de la imagen. Para preservar la información de localización de los parches, se suma una codificación posicional aprendible estándar E_{pos} a cada parche embebido, de manera que $\mathbf{x}_1^n = \mathbf{x}_1^n + E_{pos}^n$. Luego, como en un codificador transformer estándar, los parches $\{\mathbf{x}_1^n\}_{n=1}^N$ se normalizan y se proyectan a un mecanismo clásico de autoatención multi-cabeza (MSA). Particularmente, proyectamos linealmente los vectores embebidos $\{\mathbf{x}_1^n\}$ para obtener un conjunto de representaciones de consulta, clave y valor ($\{\mathbf{q}^i, \mathbf{k}^j, \mathbf{v}^j\}$). Luego, el valor se pondera mediante los pesos de atención calculados a partir de $(\{\mathbf{q}^i, \mathbf{k}^j\})$, de la forma $\mathbf{v}^j = \mathbf{v}^j \cdot \left(\text{softmax}\left(\frac{\mathbf{q}^i(\mathbf{k}^j)^T}{\sqrt{\tau}}\right)\right)$, donde τ es una variable determinada dividiendo el valor del espacio embebido de dimensión D por el número de mecanismos de autoatención k . Para la autoatención multi-cabeza, se realiza una concatenación estándar de k mecanismos de autoatención, de forma que $\text{MSA} = [\mathbf{v}^1, \mathbf{v}^2, \dots, \mathbf{v}^k]$. Esta representación se normaliza y se proyecta a un perceptrón multicapa $\mathbf{e}_1(\text{MSA})$. Este proceso se lleva a cabo a través de múltiples bloques para recuperar la caracterización embebida de los parches a lo largo del proceso de propagación. A partir de la salida del último bloque transformer, el token de clasificación (*) pasa por una cabeza MLP, codificando un embebido $\mathbf{e}_1(\text{MLP})$ que describe las características capturadas en la imagen a escala pulmonar.

Para completar el análisis, la rama local (a escala de nódulo) y la rama contextual (a escala pulmonar) se integran a partir de la representación embebida de salida. Para preservar la información de las características, se consideró una concatenación entre estos embebidos de alta densidad ($\mathbf{e}_n(\mathbf{O}) \oplus \mathbf{e}_1(\text{MLP})$). Después de esto, se realizó una proyección MLP para fusionar y potenciar la descripción contextual y local. Finalmente, se realizó la clasificación binaria (maligno/benigno).

4.3. Diseño experimental

Se utilizó el conjunto de datos LIDC-IDRI para entrenar y validar el método propuesto ²⁰. Este conjunto de datos incluye 1018 tomografías computarizadas de baja dosis, estratificadas según la severidad del cáncer basándose en el protocolo Lung-RADS. A cada nódulo se le asigna una clase de malignidad que varía del 1 al 5, representando, respectivamente, altamente benigno, moderadamente benigno, indeterminado, moderadamente maligno y altamente maligno. Para este estudio, se consideraron únicamente nódulos no indeterminados. Las clases “high-moderate-benign” y “high-moderate-malign” se agruparon en un esquema binario. El estudio se realizó con la participación de cuatro radiólogos experimentados, quienes proporcionaron diagnósticos, localización de los nódulos y características morfológicas tales como espiculación, lobulación y márgenes, entre otras. Se utilizó el valor mediano de malignidad (MMV) para asegurar el acuerdo en la anotación. En total, se incluyeron 533 nódulos en el estudio, de los cuales 174 fueron clasificados como benignos y 359 como malignos.

Debido al limitado número de datos disponibles para el entrenamiento del modelo, se optó por realizar una estrategia de validación cruzada de tipo k-fold. Esto con el propósito de evitar el sobreajuste y garantizar la capacidad de generalización del modelo. En total, el conjunto de datos se particionó en un total de 10 folds. Como optimizador del modelo, se optó por el optimizador Adam, el cual combina las ventajas de dos optimizadores populares, los cuales son el gradiente estocástico con momento (SGD) y el método de ajuste de tasas de aprendizaje adaptativas (AdaGrad)

Para la representación a escala de nódulo se consideraron parches de tamaño (32, 32, 5). Una representación intermedia RFB presenta extractores de características locales con un

²⁰ 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: vol. 38. 2. Wiley Online Library, 2011, pp. 915–931.

kernel y una dilatación de $[1, 3, 5, 7]$, respectivamente. El módulo MSA se configuró con tres cabezas para extraer patrones distintos del nódulo. El embebido resultante tiene una dimensionalidad de 32. Para la representación a escala pulmonar se consideraron 256 parches ($x_i^N = (256, 32, 32, 5)$). Además, el embebido de cada uno en un espacio de D dimensiones se fijó en 128, con un MSA que utiliza 8 cabezas. El embebido final para esta representación también tiene una dimensionalidad de 32. Se aplica una función sigmoide para la clasificación y se utiliza una función de pérdida de entropía cruzada binaria (BCE), con una tasa de aprendizaje de 0.0001.

5. EVALUACIÓN Y RESULTADOS

El enfoque propuesto se validó inicialmente mediante un estudio de ablación para medir la contribución de la rama a escala de nódulo pulmonar y de la rama a escala del pulmonar o contextual, respectivamente. La Tabla 1 resume la evaluación de los componentes, evidenciando una integración efectiva que supera a las representaciones aisladas del nódulo y del pulmón. El método propuesto alcanzó una AUC y un recall superiores, de 96.44% y 95.53%, respectivamente. El enfoque integrado presenta una ganancia de aproximadamente $\sim 2\%$ para la representación del nódulo y de hasta 10% para la información aislada a escala pulmonar.

Modelo	Acc(%)	Prec(%)	Rec(%)	AUC (%)
Escala nódulo pulmonar	88.93	90.98	92.75	94.14
Escala del pulmón	71.56	79.40	81.61	71.56
Metodo propuesto	91.90	92.68	95.53	96.44

Tabla 1. Estudio de ablación del enfoque propuesto. En este análisis se demuestra cómo la integración (ramas de nódulo y contextual) supera a las versiones aisladas.

Considerando que la integración de la información de la escala del pulmón a la información nódulo pulmonar constituye el aporte principal de este trabajo, se decidió realizar un análisis del módulo transformer, encargado de extraer la información contextual a escala de pulmón. En este estudio, se evaluó el impacto de los bloques individuales del transformer en el desempeño del modelo, variando parámetros para observar cómo esto influye en la captura de representaciones contextuales robustas que sean significativas para el diagnóstico de la clasificación del nódulo pulmonar.

Como se evidencia en la gráfica 4, al variar el número de bloques del transformer manteniendo fijo el número de cabezas de atención y el tamaño del embedding en 8 y 128 respectivamente, se demuestra que, al emplear menos de 6 bloques, se obtienen valores de AUC altamente prometedores pero inferiores al valor máximo de AUC alcanzado, lo cual sugiere que estas

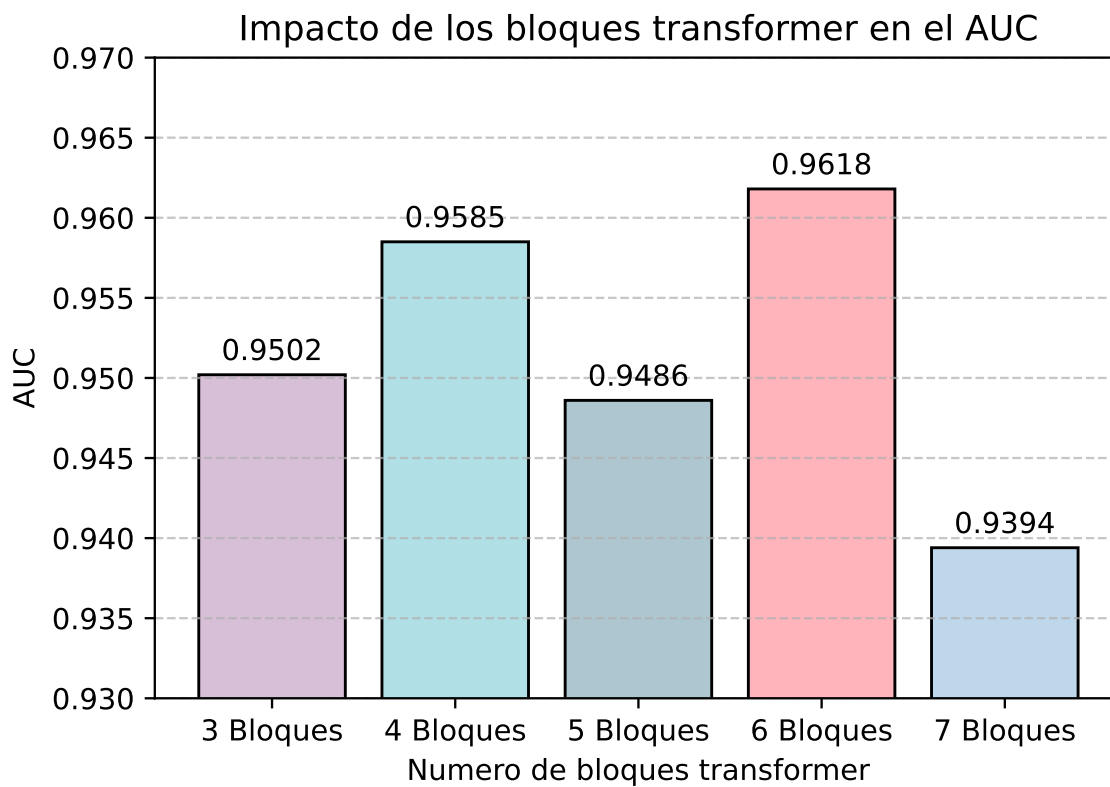


Figura 4. Analisis realizado sobre el número de bloques transformer usados para la codificación de información contextual. El estudio se realizo midiendo el rendimiento de la metrica de AUC del modelo obtenida al usar desde 3 bloques transformer hasta 7 bloques.

configuraciones podrían carecer de la profundidad necesaria para capturar de manera completa las características contextuales del pulmón. Cabe destacar que el descenso observado con 5 bloques podría atribuirse a la variabilidad estadística y estocástica inherente al modelo; es decir, aunque se observa una leve disminución en el rendimiento con 5 bloques, esta variación es moderada, puesto que el AUC obtenido al analizar únicamente la escala del nódulo pulmonar (94.14% según la Tabla 1) es considerablemente inferior. Esto evidencia que, a pesar del descenso, la integración de información contextual mediante el transformer contribuye significativamente a capturar características relevantes para la clasificación de malignidad. Por otro lado, la configuración óptima se alcanza al utilizar 6 bloques, mientras que la incorporación de 7 bloques produce un descenso marcado en el rendimiento, lo que

sugiere la presencia de sobreajuste.

Por otra parte, se realizó una comparación del enfoque propuesto con respecto a las estrategias de última generación. La Tabla 2 resume los resultados obtenidos para cada metodología presentada, de acuerdo con esquemas 3D y de atención. Como se muestra, el método propuesto supera a los métodos de última generación en precisión, recall y AUC, alcanzando 92.68%, 95.53% y 0.9644 en cada métrica respectiva. Esto demuestra que la integración en múltiples etapas mejora la caracterización contextual y local del nódulo pulmonar, logrando una mejor clasificación. Por el contrario, Al Shabi *et al.*¹² reporta una precisión de 92.81%, presentando un mayor porcentaje de nódulos correctamente predichos en comparación con otros métodos. Sin embargo, el gran desbalanceo de datos que presenta el conjunto LIDC-IDRI puede influir en esta métrica. Además, se realiza una caracterización completa del nódulo, ya que se propone una representación volumétrica. Por lo tanto, el método propuesto destaca como una estrategia volumétrica, dilatada y de atención multi-cabeza que utiliza el contexto como clave principal para clasificar nódulos pulmonares.

Modelo	Acc(%)	Prec(%)	Rec(%)	AUC (%)
3D Axial-Attention ²¹	92.81	92.59	92.36	96.17
MMVA ²²	88.76	91.82	94.63	95.19
MMEL-3DCNN ²³	90.6	-	83.7	93.9
Proposed approach	91.90	92.68	95.53	96.44

Tabla 2. Comparación de resultados frente al estado del arte.

Se llevó a cabo un análisis más detallado proyectando la representación embebida resultante en un espacio 2-dimensional de baja dimensión. Para ello, se proyectó el último embebido utilizando un Análisis de Componentes Principales (PCA) y también mediante t-SNE, recuperando dos componentes principales que representan la información de la tomografía computarizada de entrada. La Figura 5 y la Figura 6 muestran las representaciones obtenidas, destacando el poder discriminativo de los descriptores codificados, incluso siguiendo un hiperplano lineal entre las características.

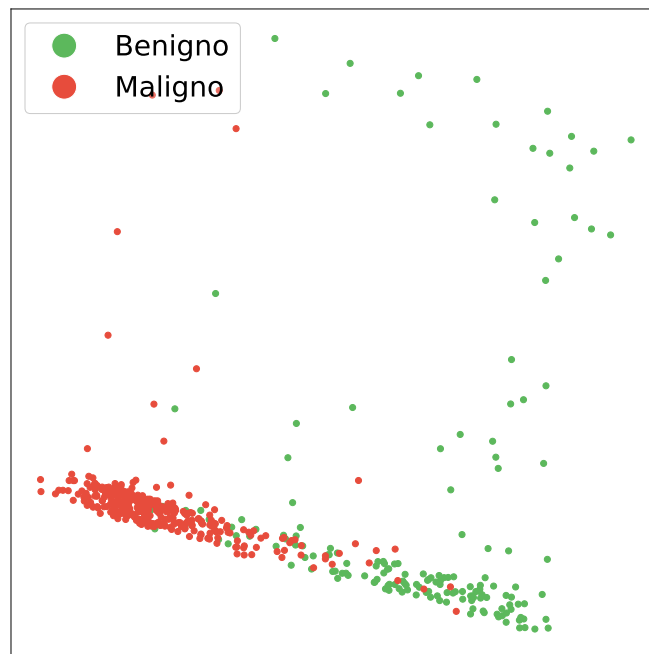


Figura 5. Análisis del embebido final que integra la información a escala de nódulo y pulmonar. La subfigura representa la reducción de dimensionalidad mediante Análisis de Componentes Principales (PCA)

En la figura 5, se presenta un análisis visual de los resultados obtenidos mediante la metodología de Análisis de Componentes Principales (PCA). Esta técnica se empleó con el fin de reducir la dimensionalidad de los embeddings generados por el modelo propuesto, permitiendo una visualización más clara y explicativa del desempeño de los resultados del modelo a la hora de discriminar entre las clases benignas y malignas de los nódulos pulmonares. La reducción dimensional con PCA facilita la observación de la distribución espacial de los datos en un espacio de menor dimensión, lo que revela de manera más directa cómo el modelo separa las clases. En esta gráfica, los nódulos malignos se representan en rojo y los benignos en verde, lo que evidencia que el modelo está llevando a cabo una adecuada clasificación, mostrando una clara separación entre las dos clases. Esta representación gráfica es coherente con los resultados obtenidos en la tabla de resultados, confirmando la efectividad del enfoque propuesto.

A continuación, en la figura 6 se utiliza t-SNE (t-Distributed Stochastic Neighbor Embedding) para proporcionar una representación aún más detallada de los resultados de clasificación del modelo. A diferencia del PCA, que es una técnica lineal de reducción de dimensionalidad, t-SNE es especialmente efectiva para preservar las relaciones locales y las estructuras no lineales entre los datos. Esto permite obtener una visualización más precisa de cómo el modelo agrupa los nódulos benignos y malignos en el espacio de características, mostrando con mayor claridad las posibles regiones de decisión del clasificador. En esta nueva visualización, los nódulos malignos se siguen representando en rojo y los benignos en verde, lo que refuerza la capacidad del modelo para diferenciar entre ambas clases, permitiendo una observación más intuitiva y destacando los patrones que podrían no ser tan evidentes mediante PCA. Esta representación gráfica complementa los resultados obtenidos con PCA, proporcionando una comprensión más profunda y detallada de la clasificación.

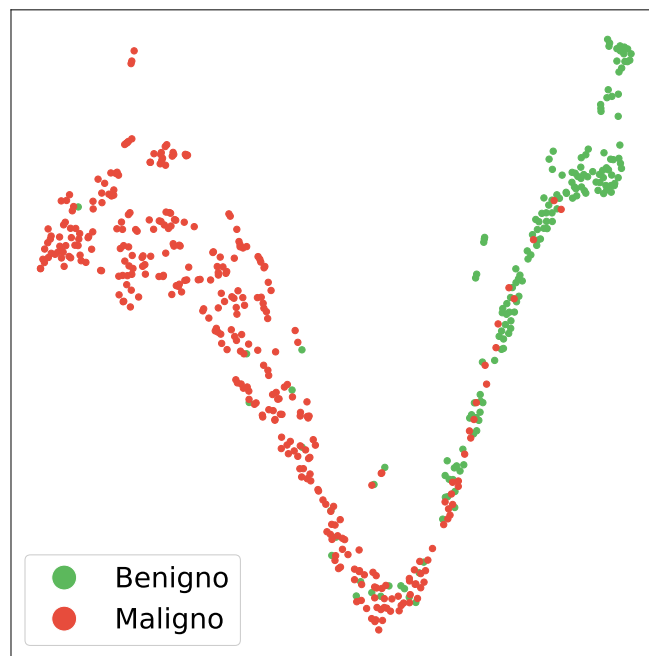


Figura 6. Análisis del embebido final que integra la información a escala de nódulo y pulmonar. La subfigura muestra los resultados del t-Distributed Stochastic Neighbor Embedding (t-SNE).

6. CONCLUSIONES Y TRABAJO FUTURO

Este trabajo introdujo un nuevo enfoque que integra un modelo de atención espacial para explotar la información del nódulo, junto con información contextual del pulmón, siguiendo una arquitectura basada en transformers. La rama a escala de nódulo captura características específicas del nódulo, incluyendo morfología, tamaño y textura, mientras que la rama a escala pulmonar proporciona un contexto global, emulando el análisis experto de TC para determinar la malignidad de un nódulo. El enfoque propuesto supera a los métodos de última generación, mostrando una dirección prometedora para incluir, como en la rutina clínica, información adicional para determinar el grado de malignidad de los nódulos. El método fue entrenado y validado utilizando el conjunto de datos públicos de cáncer de pulmón del LIDC, a partir del cual se extrajeron tanto las imágenes del slice completo del pulmón como los parches correspondientes a las zonas donde se encontraban los nódulos, gracias a la información de localización proporcionada en el dataset. Esta estrategia permitió un análisis a escala global del pulmón y a escala local del nódulo.

El método fue entrenado y validado usando el conjunto de datos públicos de cáncer de pulmón del LIDC. Donde se pudo extraer las imágenes pareadas del slice completo del pulmón, como los parches donde se encontraban los nódulos pulmonares; estos últimos pudieron obtenerse debido a que el dataset contaba con la información de la localización de cada nódulo pulmonar. Esta estrategia facilitó el análisis tanto a la escala del nódulo como a la escala global del pulmón.

Además, al emular el análisis contextual realizado por los radiólogos en las tomografías computarizadas (TC), nuestros resultados experimentales demuestran que estas características contextuales son efectivas para distinguir entre nódulos benignos y malignos. Este enfoque ayuda a mitigar la subjetividad inherente en las clasificaciones de los radiólogos, proporcionando una evaluación más consistente y reproducible. Asimismo, nuestros hallazgos destacan que incorporar información contextual en los modelos de deep learning aporta conocimientos

valiosos que mejoran la precisión en la clasificación de malignidad, reforzando la importancia de un análisis holístico en la evaluación de nódulos pulmonares.

Adicionalmente, los resultados obtenidos refuerzan la robustez del método propuesto. Con una AUC de 96.44% y un recall de 95.53%, el modelo demuestra una gran capacidad para diferenciar entre nódulos benignos y malignos. Estos indicadores sugieren que la integración de la información a escala de nódulo y pulmonar permite capturar eficazmente las características críticas para la clasificación de la malignidad, superando los enfoques convencionales. La mejora observada —aproximadamente 2% en la representación del nódulo y hasta 10% en la información contextual aislada del pulmón— evidencia que la estrategia de integración multi-etapa constituye un avance significativo en el diagnóstico asistido por computador.

Cabe destacar que realización de esta investigación se vio significativamente enriquecida gracias a los conocimientos adquiridos en las materias de maestría cursadas. La asignatura de Matemáticas Avanzadas proporcionó la base matemática necesaria para comprender arquitecturas complejas como el transformer y los mecanismos de atención, elementos esenciales para el desarrollo del presente estudio. Por otro lado, la asignatura de Seminario de Investigación I ofreció los fundamentos necesarios para llevar a cabo una investigación científica rigurosa y redactar el presente escrito, asegurando la adecuada transmisión de los resultados obtenidos. En conjunto, ambas materias aportaron conocimientos invaluable que fueron determinantes para el éxito de la presente investigación.

Como trabajo futuro, se explorará la incorporación de la geometría del pulmón en el análisis. Al considerar la estructura geométrica de las diferentes regiones pulmonares, se podrá obtener una representación más precisa de aquellas áreas que presenten indicios de malignidad. Esta integración de información geométrica permitirá identificar patrones espaciales y estructurales adicionales, lo que podría mejorar aún más la capacidad del modelo para discriminar entre nódulos benignos y malignos, y por ende, proporcionar un diagnóstico más robusto y confiable. Igualmente se realizará la exploración de conjuntos de datos alternativos para medir la capacidad de generalización, así como la comparación con referencias de la

biopsia.

BIBLIOGRAFÍA

AHMED SAIHOOD, Hossein Karshenas and NAGHSH-NILCHI, Ahmad Reza. “Multi-Orientation Local Texture Features for Guided Attention-Based Fusion in Lung Nodule Classification”. In: *IEEE Access* 11 (2023), pp. 17555–17568 (cit. on pp. 14, 24).

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: vol. 38. 2. Wiley Online Library, 2011, pp. 915–931 (cit. on p. 32).

GOULD, Michael K. et al. “Evaluation of patients with pulmonary nodules: when is it lung cancer?: ACCP evidence-based clinical practice guidelines”. In: *Chest* 132.3 (2007), 108S–130S (cit. on p. 19).

JAVED, Rabia et al. “Deep learning for lungs cancer detection: a review”. In: *Artificial Intelligence Review* 57.8 (2024), p. 197. DOI: [10.1007/s10462-023-10234-5](https://doi.org/10.1007/s10462-023-10234-5) (cit. on pp. 13, 18).

KAI XIA, et.al. “Adaptive Aggregated Attention Network for Pulmonary Nodule Classification”. In: *Applied Sciences* 11.2 (2021), p. 610 (cit. on pp. 14, 24).

LIAO, Fuyan et al. “Efficient Lung Nodule Classification Using Transferable Texture Convolutional Neural Network”. In: *IEEE Access* 8 (2020), pp. 175859–175870 (cit. on pp. 13, 24).

LIU, Hong et al. “Multi-model ensemble learning architecture based on 3D CNN for lung nodule malignancy suspiciousness classification”. In: *Journal of Digital Imaging* 33 (2020), pp. 1242–1256 (cit. on p. 36).

LIU, Yucheng et al. “Lung Nodule Malignancy Classification with Associated Pulmonary Fibrosis Using 3D Attention-Gated Convolutional Network with CT Scans”. In: *Journal of Translational Medicine* 22.1 (2024), p. 51 (cit. on p. 24).

LOVERDOS, Konstantinos et al. “Lung nodules: a comprehensive review on current approach and management”. In: *Annals of Thoracic Medicine* 14.4 (2019), pp. 226–238 (cit. on pp. 12, 18, 19).

MORENO, Alejandra et al. “A volumetric multi-head attention strategy for lung nodule classification in CT”. In: *Medical Imaging 2023: Computer-Aided Diagnosis*. Vol. 12465. SPIE, 2023, p. 1246504. DOI: [10.1117/12.2653270](https://doi.org/10.1117/12.2653270) (cit. on pp. 14, 24, 36).

MUZAMMIL, Muhammad et al. “Pulmonary Nodule Classification Using Feature and Ensemble Learning-Based Fusion Techniques”. In: *IEEE Access* 9 (2021), pp. 113415–113427 (cit. on pp. 13, 24).

PACURARI, Alina Cornelia et al. “Diagnostic accuracy of machine learning AI architectures in detection and classification of lung cancer: a systematic review”. In: *Diagnostics* 13.13 (2023), p. 2145 (cit. on pp. 13, 18).

RESEARCH ON CANCER (IARC), International Agency for. *Global Cancer Observatory*. Recuperado de IARC. 2022 (cit. on pp. 12, 18).

AL-SHABI, Mundher et al. “3D axial-attention for lung nodule classification”. In: *International Journal of Computer Assisted Radiology and Surgery* 16 (2021), pp. 1319–1324 (cit. on pp. 14, 24, 36).

AL-SHABI¹, Mundher et al. “Gated-Dilated Networks for Lung Nodule Classification in CT scans”. In: *IEEE Access* 7 (2019), pp. 178827–178838 (cit. on pp. 13, 24).

SOCIETY, American Cancer. *Detección del cáncer de pulmón*. Accedido el 25 de noviembre de 2024. 2023 (cit. on pp. 12, 18).

TEAM, National Lung Screening Trial Research. “Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening”. In: *New England Journal of Medicine* 378.24 (2017), pp. 2285–2296. DOI: [10.1056/NEJMoa1613682](https://doi.org/10.1056/NEJMoa1613682) (cit. on pp. 12, 18).

VAN RIEL, S. J. 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 pp. 13, 18).

WINER-MURAM, Helen T. “The solitary pulmonary nodule”. In: *Radiology* 239.1 (2006), pp. 34–49 (cit. on p. 19).

YI WANG, et. al. “Novel Convolutional Neural Network Architecture for Improved Pulmonary Nodule Classification on Computed Tomography”. In: *Multidimensional Systems and Signal Processing* 31 (2020), pp. 1163–1183 (cit. on pp. 13, 24).

YING CHEN, et. al. “LDNNET: Towards Robust Classification of Lung Nodule and Cancer Using Lung Dense Neural Network”. In: *IEEE Access* 9 (2021), pp. 50301–50320 (cit. on p. 24).