



EVALUACIÓN DE APROXIMACIONES DE APRENDIZAJE PROFUNDO PARA LA DETECCIÓN DE NÓDULOS PULMONARES EN TC

Manuel G. Forero, José Miguel Santos

**Universidad de Ibagué
Ibagué, Colombia**

Resumen

En 2020, según el Observatorio Mundial del Cáncer (GLOBOCAN), 1.796.144 personas murieron de cáncer de pulmón, convirtiéndolo en el cáncer con mayor tasa de mortalidad en todo el mundo. A pesar de que la tasa de supervivencia relativa a 5 años es del 63% cuando el cáncer de pulmón se detecta en un estado temprano (localizado), su diagnóstico prematuro ha sido un gran inconveniente a lo largo de la historia, principalmente porque se realiza a través de la detección de nódulos, que son difíciles de reconocer incluso para radiólogos experimentados. Por esta razón, recientemente se han propuesto retos académicos para los investigadores, donde se proporciona una base de imágenes anotadas por especialistas para desarrollar métodos más eficientes basados en aprendizaje profundo para detectar estos nódulos. Este trabajo presenta la implementación, análisis y evaluación de una técnica que emplea como entradas imágenes 3D, que se utilizan para entrenar una red neuronal convolucional VGG19, la cual fue diseñada para recibir imágenes 2D, logrando una precisión superior al 94%.

Palabras clave: cáncer pulmonar; detección de nódulos; diagnóstico temprano

Abstract

In 2020, according to the World Cancer Observatory (GLOBOCAN), 1,796,144 people died of lung cancer, making it the cancer with the highest mortality rate worldwide. Despite the fact that the 5-year relative survival rate is 63% when lung cancer is detected at an early (localized) stage, its premature diagnosis has been a major drawback throughout history, mainly because it is performed

through the detection of nodules, which are difficult to recognize even for experienced radiologists. For this reason, academic challenges have recently been proposed for researchers, where a base of images annotated by specialists is provided to develop more efficient methods based on deep learning to detect these nodules. This paper presents the implementation, analysis and evaluation of a technique that uses 3D images as inputs, which are used to train a VGG19 convolutional neural network, which was designed to receive 2D images, achieving an accuracy higher than 94%.

Keywords: lung cancer; nodule detection; early diagnosis

1. Introducción

El cáncer de pulmón es el tercer cáncer más frecuente y la principal causa de muerte relacionada con el cáncer en el mundo. Este cáncer tiene una alta letalidad con una supervivencia global del 16% a los cinco años. Los síntomas son inespecíficos, por lo que el diagnóstico suele retrasarse. Para lograr un diagnóstico más temprano e iniciar el tratamiento en una fase no avanzada del cáncer para reducir la mortalidad, se realizan tomografías computarizadas (TC) de baja dosis de radiación. Por ello, se requieren técnicas avanzadas de procesamiento de imágenes y aprendizaje automático, ya que el gran volumen de imágenes que generan los equipos médicos hace que se tenga que revisar mucha información para realizar un diagnóstico médico. Para el diagnóstico, las imágenes son analizadas por los especialistas para encontrar nódulos, medirlos y evaluarlos. Sin embargo, los nódulos encontrados en los pulmones tienen diferentes formas, dimensiones y texturas, lo que dificulta su identificación. Por este motivo, este trabajo presenta la implementación, análisis y evaluación de una técnica de Deep Learning para la detección de nódulos pulmonares en TAC, dando como resultado modelos de predicción con un alto porcentaje de precisión.

2. Materiales

Para el desarrollo de este trabajo, se realizó una búsqueda exhaustiva en bases de datos públicas de imágenes médicas de tórax para obtener la mayor cantidad de datos posible. A pesar de haber encontrado varios conjuntos de imágenes, la mayoría de estas bases de datos no contaban con las anotaciones necesarias para nuestro objetivo, y otras no tenían anotación alguna. Sin embargo, en la plataforma para el desarrollo integral de soluciones de aprendizaje automático en imagen biomédica "Grand Challenge", se propusieron dos retos similares para la detección, segmentación y reducción de falsos positivos en el estudio de nódulos pulmonares mediante técnicas avanzadas de aprendizaje automático. Allí se encuentran las siguientes bases de datos anotadas:

LNDb

Esta base de datos consta de 294 tomografías 3D de tamaño 512 x 512 x Z voxeles, en formato mhd, donde Z es el número de cortes y varía entre cada examen. Las adquisiciones se recogieron entre 2016 y 2018 en el Hospital Universitario de São João (CHUSJ) de Oporto, Portugal. Cada examen fue leído por al menos un radiólogo del CHUSJ para identificar los nódulos pulmonares y participaron un total de cinco radiólogos con al menos cuatro años de experiencia cada uno. El



proceso de anotación de los exámenes se realizó de forma ciega, es decir, los radiólogos no conocían las anotaciones de sus colegas. Los nódulos se identificaron y clasificaron en tres clases: Menos de 3 mm de tamaño, mayor o igual a 3 mm, y falsos positivos, que son lesiones que el radiólogo no considera nódulos, pero que tienen características muy similares. Las coordenadas donde se encuentran los nódulos también se incluyen en la base de datos.

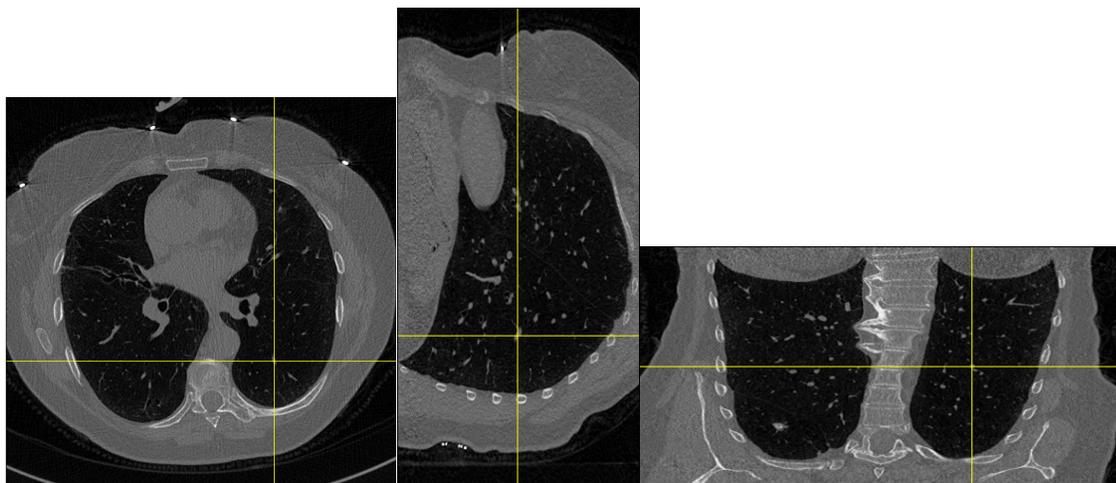


Figura 1. Corte tridimensional de un examen de la base de datos LNDb.

LUNA

La base de datos pública de este reto, LIDC/IDRI, contiene 888 tomografías 3D de tamaño $512 \times 512 \times Z$, donde Z es el número de cortes en formato mhd. Las adquisiciones sólo incluyen imágenes obtenidas con una separación entre cortes inferior a 2,5 mm y sólo se consideran verdaderos positivos los nódulos cuyo diámetro es mayor o igual a tres mm y marcados por al menos tres de los cuatro radiólogos expertos que evaluaron estas imágenes. Las anotaciones que no se ajustan a esta norma de referencia, es decir, los falsos positivos, los nódulos menores de tres mm y los nódulos señalados como positivos por sólo uno o dos radiólogos se consideran hallazgos irrelevantes. A diferencia de la base de datos LNDB, en ésta la posición del nódulo se indica únicamente mediante sus coordenadas y no mediante una máscara binaria.

Las anotaciones realizadas por los radiólogos difieren para cada reto. En LUNA, cada hallazgo se etiquetó con el ID de la tomografía, las coordenadas donde se encontraba el nódulo y su diámetro en milímetros. Se anotaron un total de 1186 hallazgos. Además, se tienen una lista de anotaciones de nódulos falsos positivos.

El desafío LNDB tiene algunas características adicionales en sus anotaciones. Cada hallazgo contiene el ID del escáner y, dado que una imagen puede contener varios nódulos, un ID único por hallazgo, sus coordenadas, el número de radiólogos que lo identificaron, su volumen y textura. Además, se tienen las anotaciones de los nódulos falsos positivos. La base de datos contiene un total de 1219 hallazgos. La textura fue cuantificada por los radiólogos y promediada en una escala de cero a cinco, en la que el cero corresponde a los hallazgos considerados como falsos positivos y aumenta progresivamente hasta el cinco a medida que los nódulos (verdaderos positivos) se vuelven más sólidos, siendo el uno la opacidad del vidrio deslustrado.



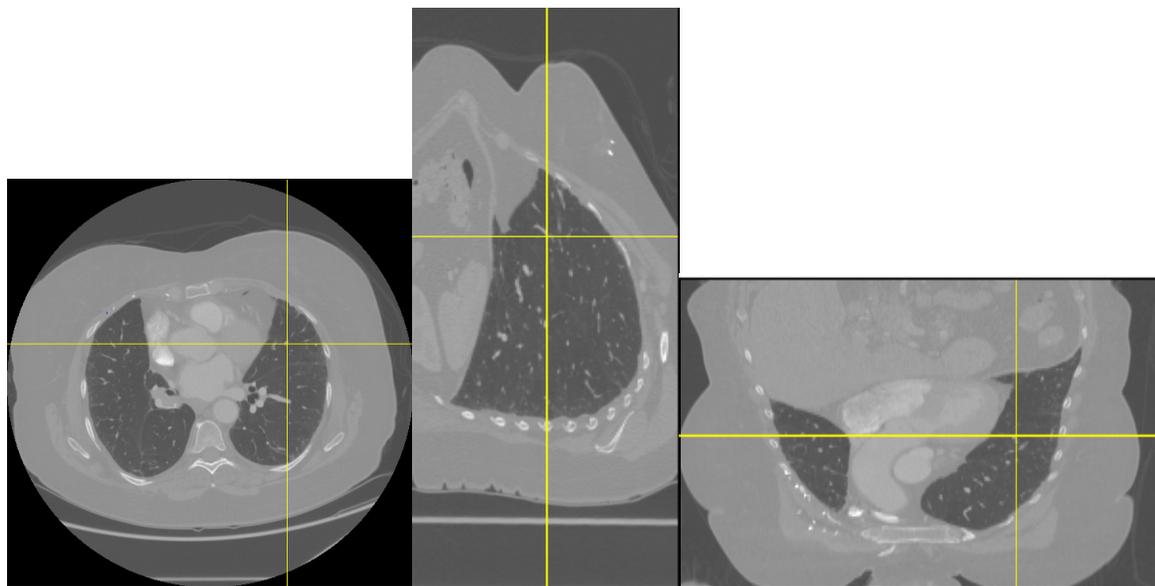


Figura 2. Corte tridimensional de un examen de la base de datos LUNA.

El método desarrollado fue escrito en el lenguaje de programación Python 3.7, utilizando las librerías SKLearn, TensorFlow2, Seaborn, Pandas, Numpy, Matplotlib y Csv y se ejecutó en un equipo Intel core i5-7200U, 8 Gb de memoria ram corriendo en el sistema operativo Windows 10. Los programas de libre acceso ImageJ e ITK-Snap fueron utilizados para el análisis de las muestras.

3. Métodos

Inicialmente con la ayuda de los programas ImageJ y ITK-Snap se realizó una exploración de los exámenes, empleando vistas ortogonales, las cuales permiten una mejor visualización de las estructuras dentro de la imagen, brindando información adicional al radiólogo para determinar si el hallazgo es positivo. Puesto que el tamaño del voxel no es constante en los distintos exámenes se reajusta el tamaño de los exámenes para que la separación entre capa y capa de la imagen fue de 1 mm, valor escogido dado que la mayoría de exámenes tenían una separación similar y, además, el diámetro de los nódulos se mide en milímetros.

Posteriormente procedemos a analizar estadísticamente las anotaciones obtenidas, encontrando datos vitales sobre el tamaño de los nódulos, su textura, posición y consenso.

Para el entrenamiento se construyó a partir de ambas de datos una base de 1956 muestras de nódulos 3D positivas y 415 negativas. Cada muestra tiene la forma de un cubo. Para determinar cuál era el mejor tamaño de los cubos, se varió su tamaño entre 10, 20 y 40 de lado, sabiendo de antemano que el máximo diámetro de los nódulos positivos encontrados en las bases de datos era de 37 mm. De tal manera que los cubos de tamaño 40 x 40 x 40 son los únicos que contienen en su totalidad los nódulos obtenidos. Se tomaron muestras de menor tamaño para observar si era posible obtener buenos resultados con prototipos más pequeños, lo cual permitiría una importante ganancia computacional.



Dos arquitecturas preentrenadas VGG16 y VGG19 fueron entrenadas. La arquitectura de la red VGG19 empleada se presenta en la Tabla 1.

Model: "VGG19"

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 40, 40, 40, 1)]	0
conv1a (Conv3D)	(None, 40, 40, 40, 64)	1792
conv1b (Conv3D)	(None, 40, 40, 40, 64)	110656
pool1 (MaxPooling3D)	(None, 20, 20, 20, 64)	0
conv2a (Conv3D)	(None, 20, 20, 20, 128)	221312
conv2b (Conv3D)	(None, 20, 20, 20, 128)	442496
pool2 (MaxPooling3D)	(None, 10, 10, 10, 128)	0
conv3a (Conv3D)	(None, 10, 10, 10, 256)	884992
conv3b (Conv3D)	(None, 10, 10, 10, 256)	1769728
conv3c (Conv3D)	(None, 10, 10, 10, 256)	1769728
conv3d (Conv3D)	(None, 10, 10, 10, 256)	1769728
pool3 (MaxPooling3D)	(None, 5, 5, 5, 256)	0
conv4a (Conv3D)	(None, 5, 5, 5, 512)	3539456
conv4b (Conv3D)	(None, 5, 5, 5, 512)	7078400
conv4c (Conv3D)	(None, 5, 5, 5, 512)	7078400
conv4d (Conv3D)	(None, 5, 5, 5, 512)	7078400
pool4 (MaxPooling3D)	(None, 3, 3, 3, 512)	0
conv5a (Conv3D)	(None, 3, 3, 3, 512)	7078400
conv5b (Conv3D)	(None, 3, 3, 3, 512)	7078400
conv5c (Conv3D)	(None, 3, 3, 3, 512)	7078400
conv5d (Conv3D)	(None, 3, 3, 3, 512)	7078400
pool5 (MaxPooling3D)	(None, 2, 2, 2, 512)	0
flatten (Flatten)	(None, 4096)	0



fc6 (Dense)	(None, 4096)	16781312
<hr/>		
dropout (Dropout)	(None, 4096)	0
<hr/>		
fc7 (Dense)	(None, 4096)	16781312
<hr/>		
dropout_1 (Dropout)	(None, 4096)	0
<hr/>		
fc8 (Dense)	(None, 4)	16388
<hr/>		
=====		
Total params: 93,637,700		
Trainable params: 93,637,700		
Non-trainable params: 0		
<hr/>		

Tabla 1. Red neuronal VGG19 empleada para determinar si un nódulo es positivo.

4. Resultados

La red VGG16 permitió obtener una precisión del 76 %, mientras que la VGG19 permitió aumentar la precisión por encima del 94%, empleando muestras constituidas por cubos de 40 mm de lado, distinguiendo los nódulos de los demás posibles hallazgos en una tomografía computacional. Estos resultados están muy por encima de los obtenidos por los especialistas. Los resultados obtenidos con los cubos más pequeños fueron más pobres obteniendo una precisión por debajo del 80%.

5. Conclusiones

El cáncer de pulmón es el más letal en el mundo y el tercero más frecuente. Aunque su diagnóstico temprano aumenta las posibilidades de supervivencia, esto es difícil y se observa en la poca coincidencia en el diagnóstico entre los especialistas. Por esta razón, es necesario desarrollar técnicas que ayuden a mejorar el diagnóstico. Para impulsar estas investigaciones se han propuesto retos que emplean imágenes anotadas por los especialistas. Como un paso más en esta área, en este trabajo se presentan los avances desarrollados en el Semillero Lún de la Universidad de Ibagué, los cuales permiten ya obtener una precisión superior al 94%. Estos resultados se obtuvieron generando nuevas bases de datos constituidas por cubos de diferentes tamaños, siendo los mejores resultados obtenidos con los cubos de lado 40, superior al radio máximo de los nódulos positivos encontrados en las bases de datos. Otras redes y nuevas arquitecturas están siendo evaluadas buscando mejorar la precisión obtenida.

Una dificultad encontrada es que las redes preentrenadas disponibles emplean únicamente como entrada imágenes 2D, por lo que se requiere el desarrollo de nuevas arquitecturas que empleen imágenes 3D.



6. Referencias

Fuentes electrónicas

- Jacobs, C., Setio, A.A.A., Traverso, A. and Ginneken, B. LUNG Nodule Analysis 2016. <https://luna16.grand-challenge.org/> (2016).
- Pedrosa, J., Ferreira, C. and Aresta, G. LNDb. Grand Challenge on automatic lung cancer patient management. <https://LNDb.grand-challenge.org> (20 November 2019).

Sobre los autores

- **Manuel Guillermo Forero Vargas:** Ingeniero Electrónico. Máster en Ingeniería Eléctrica. Máster en Imágenes Médicas y Doctor en Ingeniería Biomédica. Director del Semillero en procesamiento de imágenes y reconocimiento de patrones Lún, Decano Facultad de Ingeniería, Universidad de Ibagué. manuel.forero@unibague.edu.co.
- **José Miguel Santos Naranjo:** Estudiante de Ingeniería Electrónica, Miembro del Semillero en procesamiento de imágenes y reconocimiento de patrones Lún, Universidad de Ibagué. mitxelsk811@gmail.com.

Los puntos de vista expresados en este artículo no reflejan necesariamente la opinión de la Asociación Colombiana de Facultades de Ingeniería.

Copyright © 2021 Asociación Colombiana de Facultades de Ingeniería (ACOFI)

