

APLICACIÓN DE LAS REDES NEURONALES CONVOLUCIONALES PARA LA DETECCIÓN DEL CÁNCER DE PULMÓN

Yarens Joaquín Cruz Hernández ^a, Marcelino Rivas Santana ^a, Liliana del Carmen Valls Jorge ^b, Víctor Ferreira Moreno ^c

^a Centro de Estudio de Fabricación Avanzada y Sostenible, Facultad de Ciencias Técnicas, Universidad de Matanzas. yarens.cruz@umcc.cu, marcelino.rivas@umcc.cu

^b Departamento de Ingeniería Industrial, Facultad de Ciencias Empresariales, Universidad de Matanzas. liliana.valls@umcc.cu

^c Universidad de Ciencias Médicas de Matanzas. victorf.mtz@infomed.sld.cu

Resumen: El cáncer de pulmón es una enfermedad muy extendida a nivel mundial causando un elevado número de muertes anuales. Esta enfermedad requiere de su detección temprana para la aplicación de tratamientos efectivos que aumenten las probabilidades de supervivencia del paciente. La utilización de sistemas de diagnóstico asistido por computadora ha demostrado ser de gran ayuda en la detección de esta enfermedad, sobre todo mediante el uso de redes neuronales convolucionales para el análisis de imágenes de tomografía axial computarizada. En este trabajo se presenta una metodología para desarrollar una red neuronal convolucional atendiendo a uno de sus inconvenientes más relevantes: la selección de hiperparámetros. Esta metodología fue utilizada para crear una red neuronal convolucional que sirviera como prueba de concepto utilizando un conjunto de datos reducido. Los resultados obtenidos respaldan el uso de la metodología propuesta, al generarse una red de alta precisión, la cual al ser probada con nuevos datos clasificó correctamente el 97.5% de estos.

Palabras claves: redes neuronales convolucionales, cáncer de pulmón, diagnóstico asistido por computadora, selección de hiperparámetros.

Abstract: Lung cancer is a worldwide spread disease causing a high number of deaths every year. Early detection is a major requirement in order to apply effective treatments increasing the patient's survival probabilities. The use of computer-aided diagnose systems has proven to be very helpful for lung cancer detection, specially by means of convolutional neural networks for analyzing computer tomography images. In this work is presented a methodology for developing a convolutional neural network dealing with one of its major drawbacks: hyperparameter selection. This methodology was used to create a convolutional neural network acting as a proof-of-concept with a small dataset. The results obtained back up the use of the proposed methodology, generating a high accuracy network, which correctly classified 97.5% of the unseen test data.

Keywords: convolutional neural network, lung cancer, computer-aided diagnosis, hyperparameter selection.

1. Introducción.

El cáncer de pulmón es el más mortal de todos los tipos de cáncer. Solamente en 2018 debido a esta enfermedad se produjeron aproximadamente 1.8 millones de muertes alrededor del mundo [1]. Tan solo un 17.0 % de los pacientes diagnosticados con cáncer de pulmón en Estados Unidos sobrevive más de 5 años después del diagnóstico; mientras que en otras naciones subdesarrolladas la cifra es incluso menor [2].

En las primeras etapas de la enfermedad los tratamientos son más efectivos, por lo que su detección temprana incrementa significativamente las probabilidades de supervivencia [3]. Se ha demostrado que la disección quirúrgica ofrece grandes probabilidades de curar permanentemente al paciente en los inicios del padecimiento [4]. Sin embargo, el diagnóstico temprano del cáncer de pulmón no es una tarea fácil ya que existen muchos menos síntomas que en etapas avanzadas de la enfermedad [5].

El cáncer de pulmón es un tumor maligno caracterizado por el crecimiento descontrolado de células en el tejido pulmonar. Aunque la mayoría de los nódulos pulmonares no son cancerosos, el diagnóstico de cáncer de pulmón requiere de la identificación de dichos nódulos para poder efectuar una biopsia. Por lo tanto, la detección de nódulos adquiere una gran importancia para el diagnóstico de esta enfermedad [6]. Sin embargo, el consenso de los radiólogos en la detección de nódulos de pequeño tamaño es inferior al 52.0 % [7]. La dificultad en la detección de los nódulos está dada por la variación en la forma y el tamaño de los mismos. La Figura 1 muestra nódulos de menos de 3 mm en pacientes diagnosticados con cáncer de pulmón.

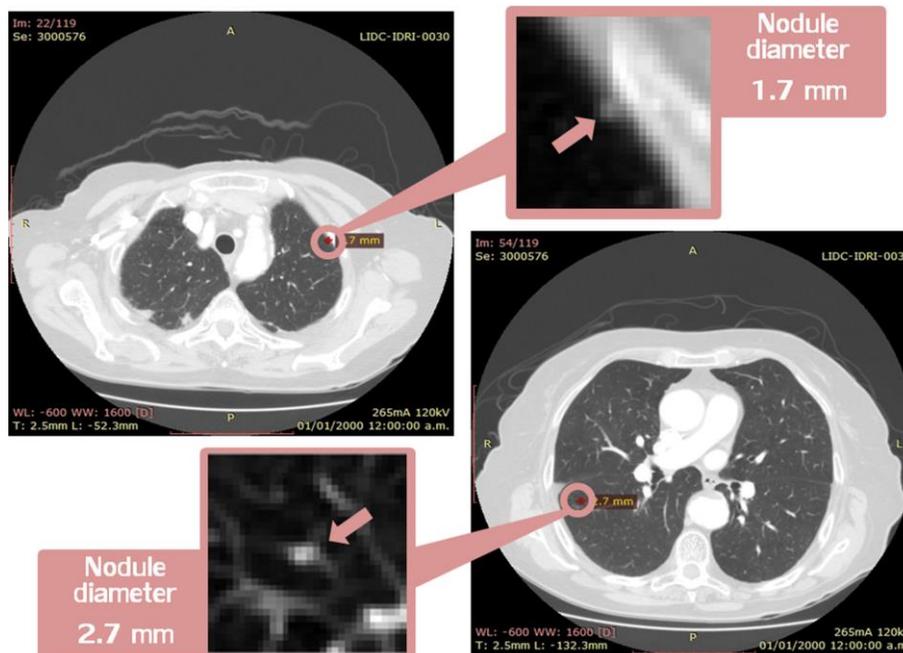


Figura 1. Imágenes de nódulos de pequeño tamaño en pacientes de cáncer de pulmón.

En las dos últimas décadas se han realizado gran cantidad de estudios sobre sistemas de Diagnóstico Asistido por Computadora (CAD, por sus siglas en inglés) aplicados a la detección de

cáncer de pulmón [8, 9]. La mayoría de estos sistemas se basan en el análisis de imágenes obtenidas mediante Tomografía Axial Computarizada (TAC). Una de las herramientas cuyo uso se ha incrementado notablemente en este ámbito son las Redes Neuronales Convolucionales (CNN, por sus siglas en inglés).

Ponnada y Srinivasu [10] desarrollaron un sistema llamado EFFI-CNN, el cual alcanzó una exactitud del 98% al analizar los conjuntos de datos de Kaggle y LIDC-IDRI. Lin et al. [11] propusieron una arquitectura de CNN y se enfocaron en optimizar, mediante el método de Taguchi, algunos hiperparámetros del modelo. Al utilizar esta red para clasificar los conjuntos de datos LIDC-IDRI y SPIE-AAPM alcanzaron exactitudes del 98.8% y 99.9% respectivamente. Otro enfoque para la aplicación de las CNN en la detección eficiente del cáncer de pulmón es la transferencia de aprendizaje, el cual fue utilizado por Almas et al. [12] partiendo de dos modelos de CNN muy conocidos: AlexNet y GoogLeNet. Por otra parte, Wang et al. [13] lograron, mediante la inclusión de bloques residuales en las CNN, una exactitud del 96.7% en el conjunto de datos Luna16. Recientemente, se han utilizado las CNN tridimensionales para la mejora de métricas como la especificidad, aprovechando la capacidad para el análisis unificado de todas las imágenes de un TAC que poseen dichos modelos [14].

Este trabajo se realiza atendiendo a la necesidad de poner los avances más recientes en el campo de la Visión por Computador en función del diagnóstico temprano del cáncer de pulmón. El objetivo principal de este trabajo es presentar la metodología utilizada para el desarrollo de un sistema de CAD de cáncer de pulmón, la cual incluye un proceso de búsqueda aleatoria de hiperparámetros. Se pretende, además, mostrar los resultados preliminares obtenidos al evaluar un conjunto de datos reducido utilizando la metodología propuesta.

2. Materiales y métodos.

Las CNN son un tipo de red neuronal artificial que implementa la operación de convolución en alguna de sus capas ocultas en lugar de la clásica multiplicación matricial [15]. Inicialmente se utilizaron para el reconocimiento de caracteres, pero adquirieron gran popularidad y versatilidad cuando se logró, mediante su uso, obtener resultados muy superiores al resto de las técnicas en el estado del arte durante el desafío de clasificación ImageNet [16]. Actualmente son uno de los modelos de Inteligencia Artificial más extendido para la clasificación de imágenes.

Las CNN se componen de varias capas sucesivas como se muestra en la Figura 1. La primera capa es la encargada de mapear cada pixel de la imagen a una neurona de la red. Para poder realizar esta operación es necesario que las dimensiones de esta capa coincidan exactamente con las de las imágenes a utilizar. En el caso que todas las imágenes no tengan la misma dimensión, estas deben ser redimensionadas.

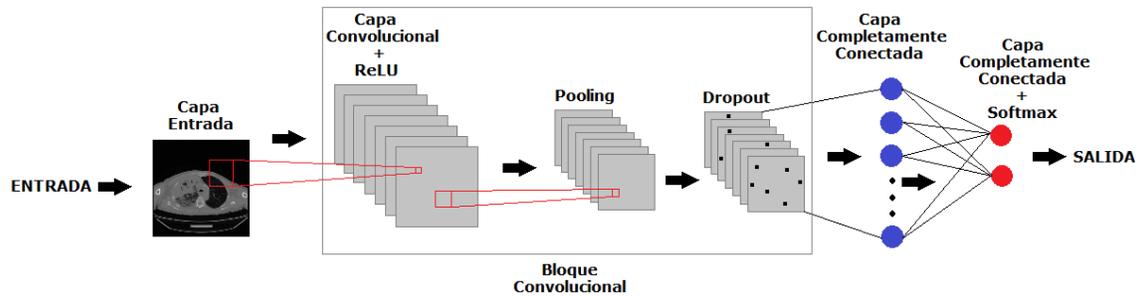


Figura 2. Estructura básica de una CNN.

Este tipo de red puede tener varios bloques convolucionales uno a continuación de otro. Cada bloque convolutivo puede estar compuesto por capas convolucionales, capas de unidades lineales rectificadas (ReLU, por sus siglas en inglés), capas de pooling y capas de dropout. Las capas convolucionales efectúan una operación de convolución en la salida de la capa precedente utilizando un conjunto de filtros que permiten distinguir las características de la imagen. La operación de convolución se describe mediante la siguiente ecuación:

$$G[m,n] = (f * h)[m,n] = \sum_j \sum_k h[j,k] f[m-j,n-k] \quad (1)$$

Donde f es la imagen de entrada mapeada en una matriz, h es un filtro y, m y n son las filas y las columnas respectivamente.

Con el objetivo de mantener la información que se propaga en la red dentro de un umbral aceptable se utilizan las capas ReLU descritas por la siguiente ecuación:

$$f(x) = \max(0, x) \quad (2)$$

Las capas de pooling permiten reducir la complejidad computacional de la red al hacer un muestreo de los resultados obtenidos en la capa anterior. Las dos estrategias de pooling más utilizadas consisten en dividir la información recibida en cuadrículas de $n \times n$ y en cada una de esas cuadrículas extraer el mayor valor o el valor promedio. Las capas de dropout permiten prevenir el sobreajuste de la red durante el entrenamiento. Esto se logra al descartar aleatoriamente, en cada época de entrenamiento, la respuesta de un cierto número de neuronas de la capa anterior.

A la salida de la CNN generalmente se utilizan capas completamente conectadas y capas softmax. Las capas completamente conectadas logran aprender las relaciones entre las diferentes características extraídas por las capas anteriores de la red al conectar todas las neuronas de la capa precedente a cada una de las neuronas contenidas en ellas. La última capa completamente conectada debe contener un número de neuronas igual a la cantidad de clases que se quiere predecir y a continuación se utiliza una capa de softmax que permite asignar valores de probabilidades de pertenencia a las clases predichas.

Las CNN son modelos de aprendizaje supervisado, es decir, requieren que durante su entrenamiento se les proporcione datos de entrada/salida. En el caso particular de este estudio, los datos de entrada son las imágenes de TAC y los de salida las categorías a las que

pertenecen (cáncer o normal), ya que inicialmente no se pretende diferenciar los diversos tipos de cáncer de pulmón. Durante el entrenamiento, la red va ajustando sus parámetros para que las características más relevantes de las imágenes tengan mayor influencia a la hora de clasificarlas. El entrenamiento de estos modelos es un proceso iterativo en el cual los datos son procesados varias veces por la red hasta que algún parámetro cumpla determinado criterio deseado o hasta que se ejecute un número de repeticiones predefinido.

En ocasiones durante el entrenamiento puede ocurrir un sobreajuste de los parámetros de la red, es decir, que memoriza las características de los datos de entrenamiento, lo que provoca un aumento en la precisión, pero disminuye la capacidad de generalización. Al ocurrir un sobreajuste, cuando se le presentan nuevos casos a la red para clasificarlos generalmente su rendimiento es deficiente. Para evitar esta situación se pueden aplicar técnicas de detención temprana del entrenamiento, para lo cual es necesario proveer a la red, además de los datos de entrenamiento, un conjunto de validación. Después de cada iteración, la red evalúa su rendimiento con los datos de validación, los cuales no influyen en el ajuste de parámetros de la misma. Mientras el rendimiento de la red continúe mejorando al validarla se sigue entrenando y si, en algún momento, el rendimiento decae se termina el entrenamiento.

Como se mencionó previamente los datos de entrada a la CNN, en este caso, son las imágenes obtenidas mediante TAC. El formato de dichos archivos al ser creados es DICOM, el cual no es soportado por softwares de propósito general. Por tanto, es necesario una etapa de preprocesamiento en la cual se conviertan los archivos a formato PNG, uno de los estándares para almacenamiento de imágenes. Por otra parte, las imágenes de TAC son creadas en escala de grises, lo que condiciona que la capa de entrada a la CNN tenga un solo canal de color. Hasta el momento, todas las imágenes de TAC que se han utilizado tienen una dimensión de 512×512 píxeles, por lo que se optó por definir la capa de entrada a la red con este mismo tamaño.

Basado en experiencias previas con este tipo de modelos se decidió inicialmente que la arquitectura de la CNN contara con 3 bloques convolucionales compuesto por una capa convolucional, una capa ReLU y una capa de pooling. Además, cada bloque convolucional podrá aleatoriamente incluir o no una capa de dropout.

Uno de los inconvenientes a tener en cuenta para el uso de las CNN radica en la selección de los hiperparámetros de red, los cuales son: número de filtros y su tamaño en las capas convolucionales, el tamaño de la cuadrícula en las capas de pooling, el número de unidades en las capas completamente conectadas y la tasa de descarte en las capas de dropout. Estos hiperparámetros tienen un impacto directo en la precisión de la red, por tanto, su correcta selección es vital para aumentar su rendimiento; sin embargo, esta tarea es complicada debido a la gran cantidad de combinaciones posibles de hiperparámetros [17].

Debido a que no existe un procedimiento analítico para determinar los hiperparámetros óptimos para una CNN, normalmente su selección se basa en un proceso de prueba y error [18]. Otros enfoques proponen la realización de una búsqueda de hiperparámetros por cuadrículas, pero en muchas ocasiones el espacio de búsqueda es tan grande que la implementación de esta solución resulta poco práctica. Para hacer una selección práctica de los hiperparámetros se puede hacer una búsqueda aleatoria, disminuyendo drásticamente el número de combinaciones [19].

Teniendo en cuenta los elementos presentados anteriormente se propone el uso de una metodología que incorpore una búsqueda aleatoria de los hiperparámetros de la CNN para el desarrollo de un sistema de CAD de cáncer de pulmón. Esta metodología se puede apreciar en la Figura 3.

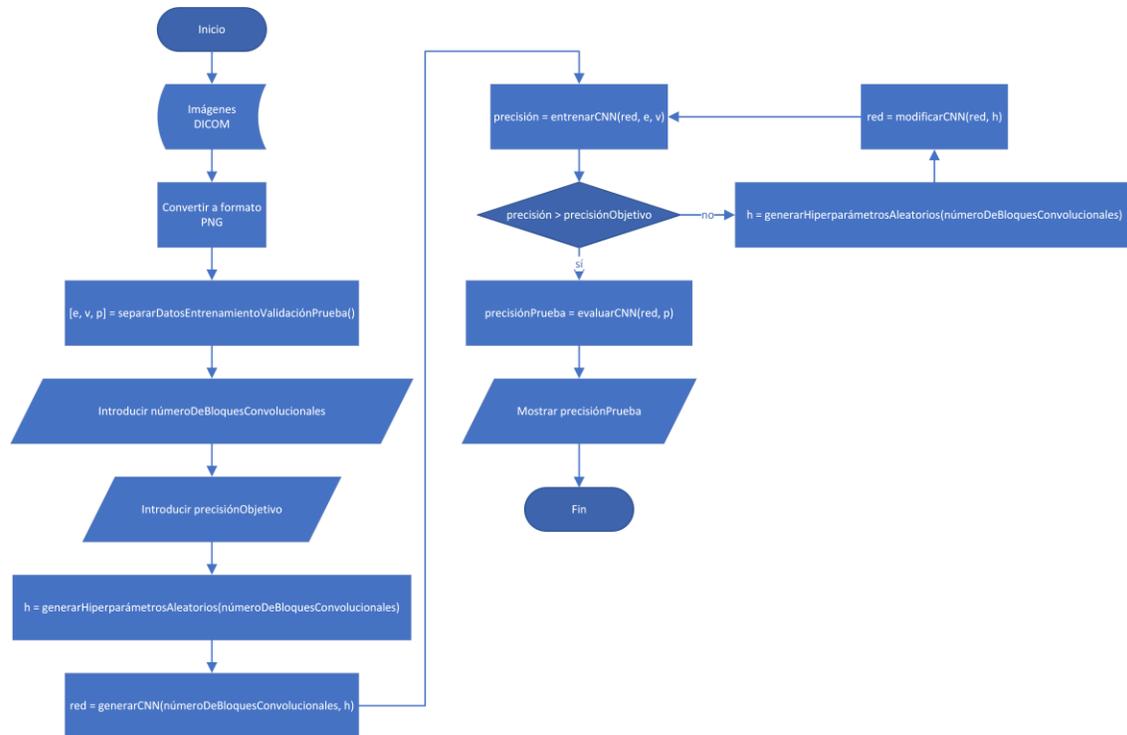


Figura 3. Metodología propuesta para la creación de la CNN, incorporando búsqueda aleatoria de hiperparámetros.

3. Análisis de los resultados.

El entrenamiento y la validación del modelo a utilizar constituyen etapas de gran importancia en el desarrollo de un sistema de CAD. Como se mencionó previamente las CNN requieren un conjunto de datos clasificados para poder llevar a cabo su entrenamiento. Mientras mayor cantidad de datos representativos de las diferentes categorías se utilicen en el entrenamiento, el modelo mejorará su capacidad de generalización a la hora de determinar a qué clase pertenece cada nuevo caso que se evalúe.

Debido a que este trabajo se basa en los resultados obtenidos en las primeras etapas del desarrollo del sistema de CAD, solo se cuenta con datos de un grupo reducido de pacientes. Los modelos se entrenaron inicialmente como una prueba de concepto y se utilizó solamente un total de 1596 imágenes. Estas fueron divididas en 798 para entrenamiento, 399 para validación y 399 para prueba. En futuras etapas del desarrollo del sistema los modelos serán reentrenados utilizando más datos, lo que permitirá mejorar su desempeño.

Para la implementación de la metodología descrita se definieron los rangos de los hiperparámetros, los cuales pueden ser vistos en la Tabla 1. Por otra parte, en la Figura 4 se

muestra la variación en la precisión de la red al ir modificándose la configuración de sus hiperparámetros en cada iteración.

Tabla 1. Rangos definidos para los hiperparámetros de la CNN.

Capa	Tamaño del filtro	Número de filtros	Tamaño de cuadrícula	Tasa de descarte	Número de unidades
Convolución 1	2-16	1-96	-	-	-
Pooling 1	-	-	2-6	-	-
Dropout 1 (opcional)	-	-	-	0-0.5	-
Convolución 2	2-14	1-48	-	-	-
Pooling 2	-	-	2-5	-	-
Dropout 2 (opcional)	-	-	-	0-0.4	-
Convolución 3	2-12	1-24	-	-	-
Pooling 3	-	-	2-4	-	-
Dropout 3 (opcional)	-	-	-	0-0.3	-
Completamente conectada	-	-	-	-	16-128
Dropout 4 (opcional)	-	-	-	0-0.5	-
Completamente conectada	-	-	-	-	2

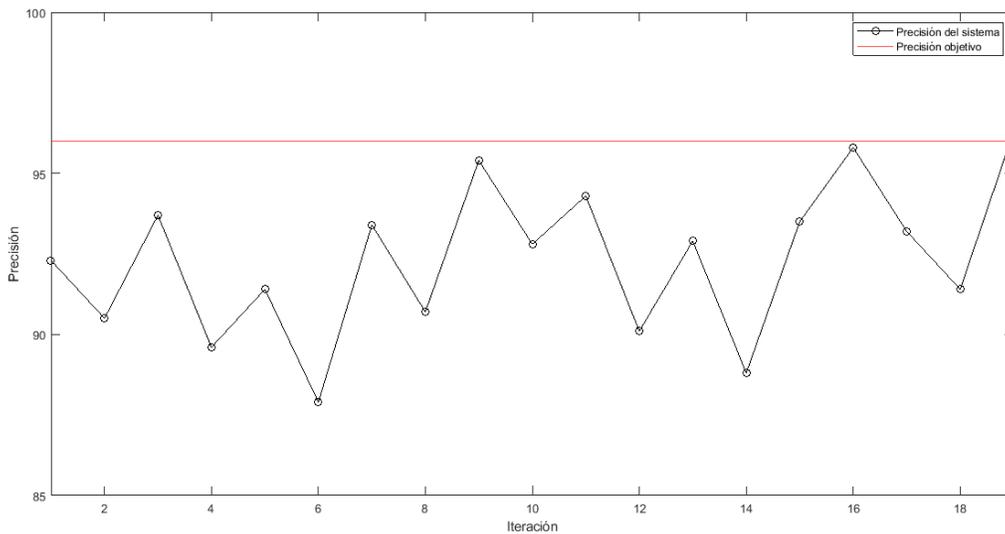


Figura 4. Precisión de la CNN al variar los hiperparámetros.

Después de 19 iteraciones se logra alcanzar una precisión superior al 96.0 %, que fue el valor definido como precisión objetivo. Es importante destacar que mediante la detención temprana se logró evitar el sobreajuste de los parámetros de la CNN cada vez que se llevó a cabo el entrenamiento, pues este se detuvo al no detectarse una disminución sostenida en el gradiente.

A continuación, se procedió a evaluar el desempeño de la red utilizándola para clasificar las imágenes de prueba, las cuales no estuvieron involucradas en el entrenamiento y la validación. Durante la prueba se obtuvo una precisión del 97.5%, con solo 6 falsos negativos al evaluar las 399 imágenes, lo que representa una tasa de falsos negativos del 1.5%, el cual es uno de los indicadores más importantes en el CAD.

4. Conclusiones.

Con el desarrollo de este trabajo se pudo verificar que la metodología presentada, la cual incluye una búsqueda aleatoria de los hiperparámetros de la red, permite desarrollar modelos de CNN que pueden ser aplicados para el CAD de cáncer de pulmón poseyendo una gran precisión. Este tipo de sistemas, no debe ser visto como el criterio definitivo para emitir un diagnóstico, sino como un apoyo al radiólogo; sin embargo, es notable la baja tasa de falsos negativos, lo que asegura una baja probabilidad de clasificar un paciente enfermo como sano y no tratar su enfermedad.

Es importante destacar que las 1596 imágenes utilizadas en este estudio corresponden solamente a 17 pacientes. Aunque esta cantidad fue suficiente para realizar una prueba de concepto a las CNN, el modelo obtenido no debe ser utilizado en la práctica hasta que no se haya reentrenado con una base de datos adecuada.

Para el entrenamiento se utilizaron imágenes de TAC de pacientes que tenían cáncer o estaban sanos, pero no se consideraron casos de pacientes que tuvieran otro tipo de afectación pulmonar. Es importante que durante el reentrenamiento del modelo se incluyan imágenes con estas características, pues uno de los mayores problemas encontrados en la literatura corresponde con el alto número de falsos positivos, debido a que se considera como cáncer cualquier tipo de anomalía, ya que no se entrenó correctamente al modelo.

Referencias.

- [1] W. H. Organization. (2018). Available: <https://www.who.int/en/news-room/fact-sheets/detail/cancer>
- [2] W. Alakwaa, M. Nassef, and A. Badr, "Lung Cancer Detection and Classification with 3D Convolutional Neural Network (3D-CNN)," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 8, pp. 409-417, 2017.
- [3] S. Sasikala, M. Bharathi, and B. R. Sowmiya, "Lung Cancer Detection and Classification Using Deep CNN," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 2, pp. 259-262, 2018.
- [4] C. Zhang, K. Dang, X. Sun, and Q. Zhou, "Toward an Expert Level of Lung Cancer Detection and Classification Using a Deep Convolutional Neural Network," *The Oncologist*, vol. 24, pp. 1-7, 2019.
- [5] W.-J. Choi and T.-S. Choi, "Automated pulmonary nodule detection system in computed tomography images: A hierarchical block classification approach," *Entropy*, vol. 15, no. 2, pp. 507-523, 2013.

- [6] G. Perez and P. Arbelaez, "Automated lung cancer diagnosis using three-dimensional convolutional neural networks," *Medical & Biological Engineering & Computing*, vol. 58, pp. 1803-1815, 2020.
- [7] S. G. A. III *et al.*, "The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Complete Reference Database of Lung Nodules on CT Scans," *Medical Physics*, vol. 38, no. 2, pp. 915-931, 2011.
- [8] F. Ciompi *et al.*, "Automatic classification of pulmonary perifissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box," *Medical Image Analysis*, vol. 26, no. 1, pp. 195-202, 2015.
- [9] S. Hawkins *et al.*, "Predicting malignant nodules from screening ct scans," *Journal of Thoracic Oncology*, vol. 11, no. 12, pp. 2120-2128, 2016.
- [10] V. T. Ponnada and S. V. N. Srinivasu, "Efficient CNN for Lung Cancer Detection," *International Journal of Recent Technology and Engineering*, vol. 8, no. 2, pp. 3499-3503, 2019.
- [11] C.-J. Lin, S.-Y. Jeng, and M.-K. Chen, "Using 2D CNN with Taguchi Parametric Optimization for Lung Cancer Recognition from CT Images," *Applied Sciences*, vol. 10, 2020.
- [12] B. Almas, K. Sathesh, and S. Rajasekaran, "A Deep Analysis of Google Net and AlexNet for Lung Cancer Detection," *International Journal of Engineering and Advanced Technology*, vol. 9, no. 2, pp. 395-399, 2019.
- [13] S. Wang, L. Dong, X. Wang, and X. Wang, "Classification of pathological types of lung cancer from CT images by deep residual neural networks with transfer learning strategy," *Open Medicine*, vol. 15, pp. 190-197, 2020.
- [14] B. Bronmans and E. Haasdijk, "Lung Nodule Segmentation Using 3D Convolutional Neural Networks," 2018.
- [15] Z. Zhang, G. Wen, and S. Chen, "Weld image deep learning-based on-line defects detection using convolutional neural networks for Al alloy in robotic arc welding," *Journal of Manufacturing Processes*, vol. 45, pp. 208-216, 2019.
- [16] D. Bacioiu, G. Melton, M. Papaelias, and R. Shaw, "Automated defect classification of SS304 TIG welding process using visible spectrum camera and machine learning," *NDT & E International*, vol. 107, 2019.
- [17] A. Baldominos, Y. Saez, and P. Isasi, "Evolutionary Design of Convolutional Neural Networks for Human Activity Recognition in Sensor-Rich Environments," *Sensors*, vol. 18, 2018.
- [18] S. Albelwi and A. Mahmood, "A Framework for Designing the Architectures of Deep Convolutional Neural Networks," *Entropy*, vol. 19, 2017.
- [19] R. Andonie and A. C. Florea, "Weighted Random Search for CNN Hyperparameter Optimization," *INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL*, vol. 15, no. 2, 2020.