

Sistema inteligente para la Detección y el Diagnóstico de Cáncer de Pulmón Mediante Ciencia de Datos

Emilio Aparicio¹, Dr. Diego H. Stalder², Dr. Jose Colbes², Dr. Jose Vazquez³

¹Alumno, Facultad de Ingeniería. Universidad Nacional de Asunción. San Lorenzo, Paraguay

²Orientador, Facultad de Ingeniería. Universidad Nacional de Asunción. San Lorenzo, Paraguay

³Orientador, Facultad Politécnica. Universidad Nacional de Asunción. San Lorenzo, Paraguay

Palabras clave:

- Cáncer de pulmón
- Detección
- YOLO
- Aumento de datos

Resumen

El cáncer de pulmón es una de las principales causas de muerte en el mundo y su detección temprana es crucial para mejorar las tasas de supervivencia. En este trabajo, se utilizó un conjunto de datos retrospectivos de imágenes de tomografía computarizada (TC) adquiridas de pacientes con sospecha de cáncer de pulmón que se sometieron a biopsias pulmonares estándar. Se llevó a cabo un análisis de los modelos de la arquitectura YOLO (You Only Look Once) para seleccionar el mejor en la detección de tumores en imágenes. Posteriormente, se realiza un análisis sobre filtros de mejora en imágenes y su impacto en las propiedades de la misma. Finalmente se llevaron a cabo pruebas de aumento de datos con el modelo seleccionado anteriormente, utilizando las imágenes con los filtros de mejora, para lograr un mejor resultado en el desempeño del modelo.

Keywords

- Lung cancer
- Detection
- YOLO
- Data augmentation

Abstract

Lung cancer is one of the leading causes of death in the world and its early detection is crucial to improve survival rates. In this work, a retrospective dataset of computed tomography (CT) images acquired from patients with suspected lung cancer who underwent standard lung biopsies was used. An analysis of the YOLO (You Only Look Once) architecture models was carried out to select the best one in detecting tumors in images. Subsequently, an analysis is carried out on image enhancement filters and their impact on their properties. Finally, data augmentation tests were carried out with the previously selected model, using the images with the enhancement filters, to achieve a better result in the performance of the model.

1. Introducción

El cáncer es la principal causa de muerte en el mundo y con tendencia a seguir creciendo en cifras año tras año debido en gran parte a las dietas basadas en comidas rápidas, poca o nula actividad física y el hábito de fumar, los cuales incrementan el riesgo de contraerlo [1]. Aproximadamente cada año se detectan 14 millones de casos nuevos y se registran 8 millones de muertes debido al cáncer en todas sus variantes, afectando a todas las regiones del mundo a pesar de los avances en el tratamiento del mismo y en su detección temprana [2].

Los tratamientos actuales para el cáncer de pulmón son cirugía, quimioterapia, radioterapia y terapia localizada. Comprender el origen y la biología del cáncer de pulmón de cada paciente es crucial para combatir la alta tasa de mortalidad de esta enfermedad; conocer los orígenes y factores ambientales y genéticos de cada paciente lleva a la utilización de tratamientos más específicos y estrategias más efectivas en cada caso. La utilización de métodos de detección modernos y rápidos ayudará a proceder al tratamiento mientras el cáncer aún se encuentre bastante localizado, lo que supone siempre una mejor respuesta a las terapias por parte del paciente [3].

Por todo lo expuesto anteriormente, este trabajo final de grado tiene como fin implementar mejoras en el aprendizaje de máquina para la detección y ayuda al diagnóstico del cáncer de pulmón utilizando la ciencia de datos basados en imágenes para lograr la detección temprana, y en el futuro añadir a la herramienta de telediagnóstico a implementar en la salud pública del Paraguay.

1.1. Objetivos del Trabajo Final de Grado

1.1.1. Objetivo General

- Analizar e implementar un sistema inteligente para la detección de tumores en pulmones utilizando aprendizaje de máquina.

1.1.2. Objetivos Específicos

- Recopilar, procesar y filtrar imágenes médicas de pacientes con cáncer de pulmón.
- Analizar modelos de detección de objetos en imágenes que permitan diagnosticar presencia de tumores en las imágenes.
- Investigar y comparar distintos métodos de mejora de imagen con el fin de mejorar el rendimiento de los modelos seleccionados.
- Aplicar métodos de aumento de datos con los métodos de mejora de imagen a los modelos seleccionados.

2. Alcance y limitaciones

En el presente trabajo se pretende implementar un sistema de detección de tumores de cáncer de pulmón en tomografías computarizadas utilizando el aprendizaje de máquina.

Inicialmente se buscarán bases de datos públicas de tomografías computarizadas que contengan tumores de cáncer de pulmón con sus respectivas etiquetas. Luego, se utilizarán estas imágenes para llevar a cabo comparaciones entre los modelos de la arquitectura YOLO.

Las métricas a utilizar para seleccionar el modelo serán la precisión, la sensibilidad y el promedio de la precisión media. Estas fueron seleccionadas debido a que el modelo se debe mejorar aumentando la cantidad de verdaderos positivos y disminuyendo la cantidad de falsos negativos, porque dado el contexto, el modelo no debe dejar de detectar tumores.

Una vez seleccionado el modelo, se procederá a implementar filtros de mejoras en imágenes y se analizarán su impacto en las mismas, esto con el fin de establecer las propiedades de la imagen que podrían tener un impacto en el experimento siguiente. Finalmente, culminado el análisis de los filtros de mejora,

se aplicará aumento de datos al modelo seleccionado, esperando mejorar su desempeño en la detección.

3. Estado del arte

La IA se puede aplicar a la medicina de diversas formas, desde el análisis de datos clínicos o genómicos, hasta el desarrollo de dispositivos médicos o robots quirúrgicos. En el ámbito de la imagenología, la IA se ha utilizado para crear algoritmos o modelos que pueden procesar, analizar y clasificar las imágenes de TC de forma automática o semiautomática, detectando y localizando los posibles tumores de pulmón, y estimando su tamaño, forma, tipo y estadio. Estos algoritmos o modelos se basan en técnicas de aprendizaje automático o aprendizaje profundo, que consisten en entrenar a la computadora con un conjunto de datos de imágenes previamente etiquetadas por expertos, para que pueda aprender a reconocer las características relevantes y generalizar a nuevos casos. En este capítulo, se revisará el estado del arte de la IA aplicada a la detección de cáncer de pulmón en las TC, describiendo los principales métodos, desafíos y resultados obtenidos hasta la fecha.

3.1. Antecedentes de detección de cáncer de pulmón

- Clasificación de tipos de cáncer de pulmón: Un equipo de investigadores de la Universidad de Texas en Dallas y el Centro Médico de la Universidad de Texas Southwestern desarrolló un modelo de CNN que podía clasificar los tipos de cáncer de pulmón (adenocarcinoma, carcinoma escamoso y carcinoma de células pequeñas) a partir de imágenes de biopsias, con una precisión del 96,7 % y una sensibilidad del 95,8 % [4].
- Estimación del estadio del cáncer de pulmón: Un equipo de investigadores de la Universidad de California en Los Ángeles y el Hospital Cedars-Sinai desarrolló un modelo de CNN que podía estimar el estadio del cáncer de pulmón (I, II, III o IV) a partir de imágenes de TC, con una precisión del 85,9 % y una sensibilidad del 83,3 % [5].
- Predicción de la supervivencia del cáncer de pulmón: Un equipo de investigadores de la Universidad de Toronto y el Hospital Princess Margaret desarrolló un modelo de CNN que podía predecir la supervivencia del cáncer de pulmón a partir de imágenes de TC, con un coeficiente de concordancia de 0,66 y un error cuadrático medio de 0,12 [6].
- Segmentación de tumores de pulmón: Un equipo de investigadores de la Universidad de Ciencia y Tecnología de China y el Hospital de la Universidad de Anhui desarrolló un modelo de CNN que podía segmentar los tumores de pulmón a partir

de imágenes de TC, con un coeficiente de Dice de 0,89 y una precisión del 98,7 % [7].

3.2. Sensibilidad mínima

Según una revisión sistemática publicada en el año 2019, el estándar de sensibilidad mínima que debe tener un modelo de detección de tumores de cáncer de pulmón para que sea aceptado como fiable es del 80 % [8]. Esto significa que el modelo debe ser capaz de identificar correctamente al menos el 80 % de los casos verdaderos de cáncer de pulmón. La sensibilidad es una medida de la capacidad de un modelo para evitar los falsos negativos, es decir, los casos en los que el modelo no detecta el cáncer cuando este está presente.

4. Marco Teórico

4.1. Cáncer de Pulmón

El cáncer de pulmón es un tipo de cáncer que se origina en las células del tejido pulmonar, principalmente en las que recubren los bronquios y los bronquiolos. Es la principal causa de muerte por cáncer en el mundo, tanto en hombres como en mujeres, y se estima que en el año 2020 se produjeron 2,2 millones de casos nuevos y 1,8 millones de fallecimientos por esta enfermedad [9]. El cáncer de pulmón se clasifica en dos tipos principales según el aspecto de las células tumorales al microscopio: cáncer de pulmón de células no pequeñas (CPCNP) y cáncer de pulmón de células pequeñas (CPCP) [10].

4.1.1. Detección temprana del Cáncer de Pulmón

La detección temprana del cáncer de pulmón es una estrategia que busca identificar la enfermedad en una etapa inicial, cuando el tratamiento tiene más probabilidades de ser efectivo y curativo. La detección temprana del cáncer de pulmón puede realizarse mediante pruebas de cribado o tamizaje, como la tomografía computarizada (TC) de baja radiación, o mediante el diagnóstico oportuno, que implica la reducción del tiempo entre la aparición de los primeros síntomas y el inicio del tratamiento.

4.2. Tomografía Computarizada

La tomografía computarizada (TC) es una técnica de imagen médica que permite obtener imágenes tridimensionales de la estructura interna de un cuerpo, a partir de múltiples proyecciones bidimensionales obtenidas con rayos X [11]. La TC se basa en el principio de que los tejidos del cuerpo humano tienen diferentes coeficientes de atenuación de los rayos X, es decir, que absorben o dispersan la radiación de forma distinta según su densidad y composición [12]. La TC utiliza una fuente de rayos X que gira alrededor del

paciente, emitiendo un haz de radiación que atraviesa el cuerpo y es captado por un detector situado en el lado opuesto. El detector mide la intensidad de la radiación que llega, que depende de la atenuación que ha sufrido el haz al pasar por los tejidos. El proceso se repite desde diferentes ángulos, generando una serie de proyecciones que contienen la información de la distribución de la atenuación en el plano perpendicular al haz [11].

4.3. Inteligencia Artificial

La inteligencia artificial (IA) es la ciencia y la ingeniería de crear máquinas inteligentes que imitan las capacidades de resolución de problemas y toma de decisiones de la mente humana. La IA se puede clasificar según el nivel de inteligencia que exhibe, el tipo de aprendizaje que utiliza y el dominio de aplicación que aborda.

4.4. Redes Neuronales

Las redes neuronales son algoritmos de inteligencia artificial que imitan el funcionamiento del cerebro humano para procesar datos y resolver problemas [13]. Están formadas por capas de unidades de procesamiento llamadas neuronas artificiales, que se conectan entre sí mediante pesos sinápticos [14]. Cada neurona recibe una entrada, la multiplica por su peso, le suma un sesgo y la pasa por una función de activación para producir una salida [15]. La salida de una neurona se convierte en la entrada de la siguiente, creando una red de propagación de información.

4.5. Redes Convolucionales

Las redes convolucionales son un tipo especial de redes neuronales que se inspiran en el funcionamiento de la corteza visual del cerebro para procesar datos tridimensionales, como imágenes, voz o audio [16]. Están formadas por diferentes tipos de capas, como las capas de convolución, agrupamiento y conectada, que realizan diferentes funciones según su posición en la red [14]. El objetivo de las redes convolucionales es extraer las características relevantes de los datos de entrada, mediante la aplicación de filtros o kernels, y generar el resultado deseado, que puede ser una clasificación, una segmentación o una reconstrucción de la entrada [15].

4.5.1. Capa de Convolución

La capa de convolución es una de las capas principales de la red neuronal convolucional, que se inspira en el funcionamiento de la corteza visual del cerebro para procesar datos tridimensionales, como imágenes, voz o audio [16]. La capa de convolución aplica un filtro o kernel a una región de la entrada, obteniendo una característica o mapa de características.

El filtro se desliza sobre la entrada, multiplicando sus valores por los del filtro y sumándolos, creando así una matriz de salida. El filtro se puede interpretar como un detector de patrones, que resalta las partes de la entrada que son similares al filtro. La capa de convolución puede tener varios filtros, cada uno con un tamaño, una forma y unos valores diferentes, que extraen diferentes características de la entrada [14] [17].

4.5.2. Capa de Agrupamiento

La capa de agrupamiento es otra de las capas principales de la red neuronal convolucional, que se utiliza para reducir la dimensionalidad y el ruido de los mapas de características generados por la capa de convolución [16]. La capa de agrupamiento aplica una operación de reducción, como el máximo, el mínimo o el promedio, a una región de la entrada, obteniendo un valor resumen. El objetivo de la capa de agrupamiento es simplificar la información, eliminando los detalles innecesarios y conservando las características más importantes. La capa de agrupamiento también mejora la invariancia de la red, es decir, su capacidad de reconocer los objetos independientemente de su posición, tamaño u orientación [14] [15].

4.5.3. Capa Conectada

La capa conectada es la última capa de la red neuronal convolucional, que se encarga de generar el resultado final de la red, que puede ser una clasificación, una segmentación o una reconstrucción de la entrada [16]. La capa conectada conecta todas las neuronas de la capa anterior con las de la capa siguiente, realizando una operación de multiplicación de matrices, seguida de una función de activación. La capa conectada se puede interpretar como un clasificador o un generador, que utiliza las características extraídas por las capas anteriores para producir la salida deseada [14] [17].

5. Metodología y resultados

5.1. Base de datos y materiales

5.1.1. Base de datos

Para llevar a cabo los experimentos, se utilizó un conjunto de datos retrospectivo que consiste en imágenes DICOM de TC de pacientes con sospecha de cáncer de pulmón que se sometieron a biopsias pulmonares estándar [18]. Las etiquetas se generaron mediante la colaboración de cinco radiólogos académicos especializados en cáncer de pulmón. Dos de los radiólogos tenían más de 15 años de experiencia, mientras que los otros tres tenían más de 5 años de experiencia en el campo.

La selección de la muestra se realizó a partir de un conjunto de 1886 imágenes tomográficas. Estas

imágenes se distribuyeron en tres conjuntos: entrenamiento, validación y prueba de tal forma evaluar el desempeño del modelo en diferentes etapas del proceso de entrenamiento y asegurar la generalización de los resultados.

Tabla 1: Distribución de las imágenes de la base de datos para los experimentos

Tarea	Cantidad	Proporción (%)
Entrenamiento	721	38.23
Validación	309	16.38
Test	856	45.39

5.1.2. Entorno de ejecución

■ Computadora Personal

La computadora fue adquirida con el propósito específico de servir como una herramienta de apoyo para las actividades académicas relacionadas con mi formación universitaria. Esta elección se basó en la necesidad de contar con un ambiente de trabajo dedicado y eficiente, capaz de proporcionar la potencia y flexibilidad requeridas para llevar a cabo investigaciones y tareas académicas de alta complejidad, en consonancia con los estándares y requisitos de mi programa de estudios. Cuenta con las siguientes especificaciones:

- Modelo: Acer Nitro 5 AN515-54-728C
- Procesador: Intel i7-9750H de 6 núcleos y 12 subprocesos, con una frecuencia base de 2.6GHz
- Memoria RAM: 16 GB - DDR4
- Tarjeta gráfica: GPU NVIDIA GeForce RTX 2060
- Sistema operativo: Windows 10, Python 3.8.5 y PyTorch 2.0

5.2. Descripción de experimentos

Para lograr nuestro objetivo de optimizar la detección de tumores en imágenes, hemos dividido el proceso experimental en cuatro partes distintas. Cada una de estas etapas cumple un papel esencial en nuestra investigación:

- Experimento 1 - Elección del Modelo: En la primera parte, se realizará la selección del modelo de detección de tumores. Esta elección se basará en parámetros específicos que se consideren más adecuados para el contexto de nuestro estudio.
- Experimento 2 - Modificaciones de Imágenes: En la segunda parte, se aplicarán métodos de modificación de imágenes a toda la base de datos.

El propósito de esta etapa es modificar las imágenes de tal manera que sus nuevas características puedan tener un impacto positivo en el entrenamiento del modelo. Además, se llevará a cabo una verificación estadística utilizando el método de Wilcoxon para confirmar que las bases de datos resultantes sean estadísticamente diferentes entre sí.

- Experimento 3 - Entrenamiento del Modelo: En la tercera parte, el modelo seleccionado en el primer paso se entrenará con las imágenes modificadas. El objetivo aquí es analizar cualquier mejora en la detección de tumores, si es que existe, y determinar cómo las modificaciones específicas de las imágenes influyen en esta mejora.
- Experimento 4 - Aumento de Datos en Línea: La cuarta y última parte del experimento implica la aplicación de un método de aumento de datos en línea. Aquí, se aplicarán diversos filtros con la expectativa de mejorar el rendimiento del modelo, contribuyendo aún más a la optimización de la detección de tumores en imágenes.

5.3. Arquitectura del modelo propuesto

Para llevar a cabo el experimento, se utilizaron las versiones más recientes del sistema YOLO (You Only Look Once). La arquitectura YOLO, en general, se caracteriza por su enfoque You Only Look Once en el cual realiza una única pasada a través de la imagen completa para identificar objetos y generar predicciones de sus ubicaciones. Esto permite una detección más rápida en comparación con enfoques basados en regiones. Además, YOLO utiliza una red neuronal convolucional (CNN) para extraer características relevantes de la imagen y un módulo de detección final que combina información espacial y de características para generar las predicciones de detección.

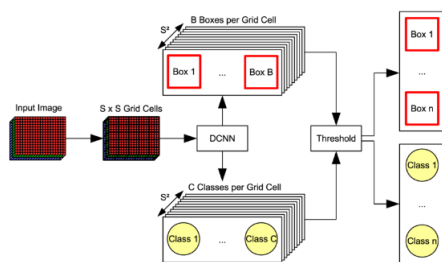


Figura 1: Resumen de la arquitectura YOLO
Fuente: [19]

5.4. Algoritmos de mejora en imágenes

Para mejorar el rendimiento de estos modelos, se acostumbra aplicar algoritmos de mejora de distintas características de una imagen, tales como el contraste, el brillo, el ruido, la resolución, la orientación, el

color, la textura, etc. Estos algoritmos buscan optimizar la calidad de la imagen y facilitar la extracción de información relevante por parte de los modelos de detección de objetos.

Un ejemplo de éxito de la aplicación de algoritmos de mejora de características de una imagen es el trabajo de Zhang et al. (2020) [20], que propusieron un método basado en el aprendizaje profundo para mejorar el contraste y el brillo de las imágenes de rayos X de tórax, con el fin de mejorar la detección de neumonía causada por COVID-19. Los autores demostraron que su método mejoraba significativamente la precisión y el tiempo de ejecución de los modelos de detección de objetos en comparación con otros métodos existentes.

5.4.1. Filtro Anisotrópico

Este método reduce el efecto de aliasing, que es la aparición de bordes dentados o pixelados en las imágenes, y conserva los detalles y la nitidez de las texturas [21]. El filtro anisotrópico requiere un mayor procesamiento y memoria que otros métodos de filtrado, como el bilineal o el trilineal, pero ofrece mejores resultados visuales [22].

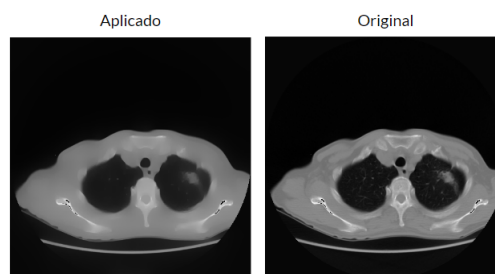


Figura 2: Demostración de aplicación del filtro anisotrópico

Fuente: Elaboración propia

5.4.2. Filtro BBHE

Este método mejora el contraste de las imágenes, que es la diferencia entre las regiones claras y oscuras, y preserva los detalles y la luminosidad de las texturas. El filtro BBHE requiere un menor procesamiento y memoria que otros métodos de mejora de contraste, como el CLAHE o el DSIHE, pero ofrece resultados visuales satisfactorios [23].

5.4.3. Filtro Bilateral

Es un método que alisa las imágenes sin modificar los bordes, que son las transiciones bruscas entre las regiones claras y oscuras, y disminuye el ruido, que son las variaciones aleatorias de intensidad.

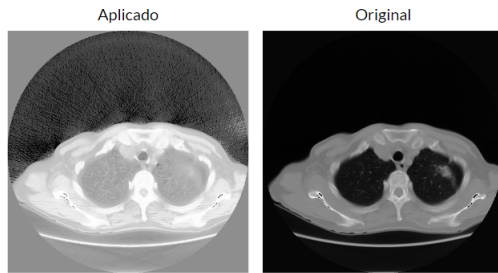


Figura 3: Demostración de aplicación del filtro BBHE
Fuente: Elaboración propia

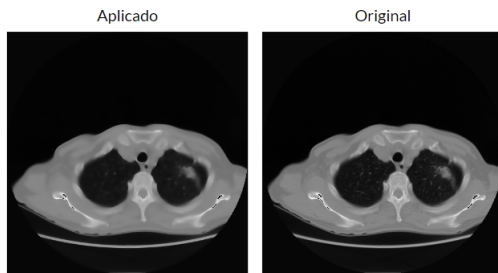


Figura 4: Demostración de aplicación del filtro bilateral
Fuente: Elaboración propia

5.4.4. Filtro CLAHE

Es un método que mejora el contraste de las imágenes, que es la diferencia entre las regiones claras y oscuras, y evita la sobre-amplificación del ruido, que son las variaciones aleatorias de intensidad.

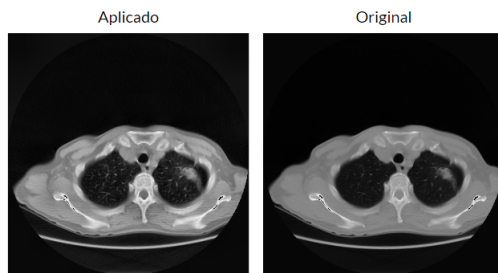


Figura 5: Demostración de aplicación del filtro CLAHE
Fuente: Elaboración propia

5.4.5. Filtro Morfológico

El filtro morfológico es un conjunto de operaciones que sirven para procesar imágenes basadas en formas geométricas. En un filtro morfológico, cada píxel de la imagen se ajusta en función del valor de otros píxeles de su entorno, que se definen mediante un elemento estructural.

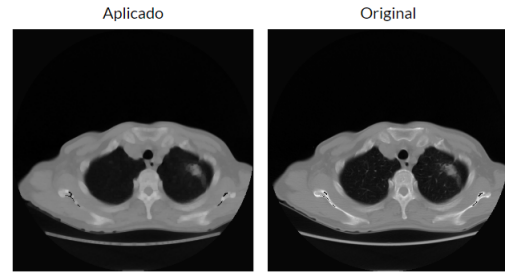


Figura 6: Demostración de aplicación del filtro morfológico
Fuente: Elaboración propia

5.4.6. Filtro Stretch Contrast

Es una técnica de realce de imagen que intenta mejorar el contraste al estirar los valores de intensidad de una imagen para llenar el rango dinámico completo. La función de transformación utilizada es siempre lineal y monótonamente creciente.

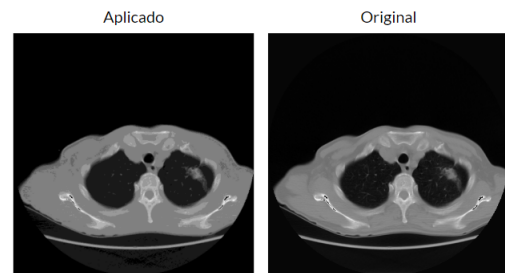


Figura 7: Demostración de aplicación del filtro stretch contrast
Fuente: Elaboración propia

5.4.7. Filtro Suavizado

El suavizado es una operación de generalización que elimina los ángulos agudos en las líneas o los contornos. A menudo se utiliza el suavizado para mejorar la estética del resultado cartográfico.

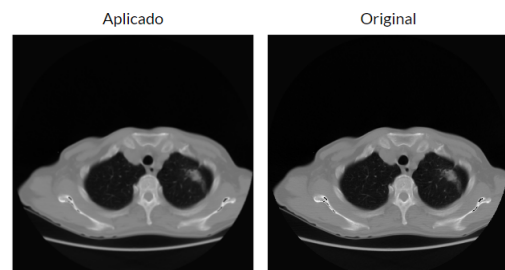


Figura 8: Demostración de aplicación del filtro suavizado
Fuente: Elaboración propia

5.5. Métricas en imágenes

5.5.1. Error de brillo medio absoluto

Se define como la diferencia absoluta entre el valor medio del brillo de la imagen de entrada y la imagen de salida. Es decir, $AMBE = |MB(X) - MB(Y)|$, donde X e Y son la imagen de entrada y la imagen de salida, respectivamente, y $MB(.)$ es el proceso de calcular el brillo medio. Un valor menor de $AMBE$ indica que el brillo se conserva mejor [24].

5.5.2. Entropía

La entropía de una imagen es una medida de la aleatoriedad o la complejidad de la información que contiene. Se puede definir como la suma de las probabilidades de cada nivel de gris multiplicadas por el logaritmo en base 2 de dichas probabilidades, con signo negativo. Es decir, $E = -\sum(p_i \log_2(p_i))$, donde p_i es la probabilidad de que un píxel tenga el nivel de gris i , y la suma se hace sobre todos los niveles de gris posibles (256 para imágenes de 8 bits). Un valor mayor de entropía indica que la imagen tiene más variación o detalle, y un valor menor indica que la imagen es más uniforme o borrosa [25].

5.5.3. Medida del índice de similaridad estructural

La medida del índice de similaridad estructural (SSIM, por sus siglas en inglés) es un método para predecir la calidad percibida de imágenes y videos digitales, así como para medir la similitud entre dos imágenes. El índice SSIM es una métrica de referencia completa; es decir, la medición o predicción de la calidad de imagen se basa en una imagen inicial sin compresión o sin distorsión como referencia. El SSIM se basa en la hipótesis de que el sistema visual humano está muy adaptado para extraer información estructural de las imágenes, y que una medida de la degradación de la información estructural puede dar una buena aproximación de la calidad de imagen percibida [26].

5.5.4. Relación pico de señal a ruido

La relación pico de señal a ruido (PSNR, por sus siglas en inglés) es un término de ingeniería que define la relación entre la potencia máxima posible de una señal y la potencia del ruido que afecta a su representación fidedigna. Debido a que muchas señales tienen un rango dinámico muy amplio, el PSNR se expresa usualmente como una cantidad logarítmica usando la escala de decibelios. El PSNR se utiliza comúnmente para medir la calidad de reconstrucción de imágenes y videos sujetos a compresión con pérdidas [27].

5.5.5. Índice de conservación de contornos

La medida del índice de conservación de contornos (EPI, por sus siglas en inglés) es un método para evaluar la capacidad de los filtros de imágenes SAR para preservar los bordes y las características de las imágenes originales. El índice de conservación de contornos se basa en la idea de que los bordes y las características de una imagen son importantes para la interpretación y el análisis de la imagen, y que un buen filtro debe reducir el ruido sin degradar la información estructural de la imagen. La diferencia con otras técnicas como el PSNR o el SSIM es que estos métodos se centran en la fidelidad global de la imagen, mientras que el EPI se enfoca en la preservación local de los bordes y las características [28].

5.5.6. Contraste

El contraste de una imagen es una medida de la diferencia entre las zonas claras y oscuras de la imagen. Una imagen con alto contraste tiene una gran diferencia entre las zonas claras y oscuras, lo que hace que los detalles sean más visibles y la imagen más nítida. Una imagen con bajo contraste tiene una pequeña diferencia entre las zonas claras y oscuras, lo que hace que los detalles sean menos visibles y la imagen más borrosa. El contraste de una imagen se puede modificar para mejorar la calidad visual de la imagen o para resaltar ciertas características de interés [29].

5.5.7. Razón de mejora de contraste

La razón de mejora de contraste (CIR, por sus siglas en inglés) es un método para evaluar el efecto de las técnicas de mejora de contraste en las imágenes digitales. La razón de mejora de contraste se define como la relación entre el contraste de la imagen mejorada y el contraste de la imagen original. Un valor mayor de CIR indica una mayor mejora de contraste, mientras que un valor menor indica una menor mejora de contraste [30].

5.6. Pruebas y Resultados

5.6.1. Entrenamientos iniciales de modelos YOLOV7 y YOLOV8

Esta etapa tiene como propósito elegir el modelo que logre los mejores resultados en la identificación y localización de tumores en las imágenes que conforman el conjunto de datos que se emplea como base. Para hacer esta selección, se utilizan las métricas que se han descrito anteriormente, que son los criterios que permiten medir el rendimiento y la precisión de los modelos. El modelo que presente los valores más altos en estas métricas será el escogido como el más óptimo para la detección de tumores. Los resultados se muestran en la tabla 2.

Tabla 2: Resumen de entrenamientos

Modelo	Precisión	Sensibilidad	mAP@0.5	mAP@0.5:0.95
YOLOv7	0.752	0.724	0.75	0.321
YOLOv8 - Nano	0.913	0.857	0.908	0.438
YOLOv8 - Small	0.884	0.842	0.905	0.452
YOLOv8 - Medium	0.869	0.779	0.861	0.401
YOLOv8 - Large	0.893	0.806	0.881	0.408

En la Tabla 2, se puede observar que el modelo Nano obtuvo los mejores resultados en 3 de los 4 parámetros utilizados para evaluar el rendimiento del entrenamiento. Esto indica que el modelo Nano logró una alta precisión, sensibilidad y mAP@0.5, lo que implica una detección precisa y confiable de los tumores pulmonares en las imágenes de tomografía computarizada. Es importante destacar que, dado que en este caso solo se está considerando una clase de detección, que es el tumor, es probable que este sea el factor que influye en el rendimiento superior del modelo Nano en comparación con los otros modelos YOLOv8 (Small, Medium y Large).

5.6.2. Cálculos de Métricas en Imágenes

Las métricas se utilizan con el fin de comprender el impacto en los diversos parámetros de la imagen. Esto permite, en caso de observar mejoras en algún modelo, identificar con precisión los factores responsables de estos avances. Este enfoque ayuda a establecer conexiones directas entre los cambios en las métricas y las posibles mejoras en el desempeño de los modelos, proporcionando una comprensión detallada de las causas detrás de cualquier mejora observada. Los resultados de las métricas se muestran en la tabla 3.

Tabla 3: Resultados de cada métrica para cada método de mejora de imagen

Método	AMBE	ENTROPY	SSIM	PSNR	EPI	CONTRAST	CIR
CLAHE	14,546	6,701	0,728	22,851	1,568	64,740	1,120
ANISOT.	12,614	6,411	0,649	23,452	0,351	49,430	0,850
BBHE	74,662	5,538	0,377	9,979	2,165	72,387	1,259
BIL.	4,389	5,544	0,833	30,714	0,579	56,182	0,969
MORFOL.	6,356	5,043	0,733	25,463	0,481	53,372	0,922
STRETCH	49,880	2,165	0,154	10,709	0,041	2,342	0,041
SUAVI.	4,283	5,608	0,872	30,957	0,677	55,794	0,964

De estos resultados obtenidos, podemos obtener la siguiente información:

- En la métrica AMBE, el que obtuvo mejor desempeño es el Filtro de Suavizado, ya que es el que obtuvo menor valor. También se deben tener en cuenta el Filtro Bilateral y el Morfológico, ya que obtuvieron valores muy cercanos.
- En la Entropía, el que obtuvo mejor desempeño es el CLAHE, ya que obtuvo el mayor valor. También se debe tener en cuenta el Filtro Anisotrópico, que obtuvo un valor cercano.
- En la métrica SSIM, el que obtuvo mejor desempeño es el Filtro de Suavizado, ya que es el que obtuvo mayor valor. También se debe tener en

cuenta el Filtro Bilateral, que obtuvo un valor cercano.

- En la métrica PSNR, el que obtuvo mejor desempeño es el Filtro de Suavizado, ya que es el que obtuvo mayor valor. También se debe tener en cuenta el Filtro Suavizado, que obtuvo un valor cercano.
- En la métrica EPI, el que obtuvo mejor desempeño es el BBHE, ya que obtuvo el mayor valor. También se debe tener en cuenta el CLAHE, que obtuvo un valor cercano.
- En el Contraste, el que obtuvo mejor desempeño es el BBHE, ya que obtuvo el mayor valor. También se debe tener en cuenta el CLAHE, que obtuvo un valor cercano.
- En el CIR, el que obtuvo mejor desempeño es el BBHE, ya que obtuvo el mayor valor. También se debe tener en cuenta el CLAHE, que obtuvo un valor cercano.

5.6.3. Aplicación del Método de Wilcoxon de a Pares

Con base en los resultados obtenidos en la fase anterior, se procede a realizar comparaciones entre el conjunto de valores obtenidos de un método particular y los valores obtenidos de los otros métodos disponibles. El método que haya mostrado el mejor desempeño se elige como punto de comparación primario. El objetivo de este análisis es detectar diferencias estadísticamente significativas entre los datos generados por cada par de métodos comparados.

Cuando se hace una comparación estadística, se quiere saber si hay una diferencia real entre los grupos de datos que se analizan. El valor de p es una medida que indica la probabilidad de obtener el resultado observado o uno más extremo, asumiendo que no hay ninguna diferencia entre los grupos. Un valor de p cercano a cero significa que es muy improbable que el resultado se deba al azar, y por lo tanto, hay una diferencia significativa entre los grupos. Un valor de p distinto de cero significa que es posible que el resultado se explique por la variabilidad de los datos, y por lo tanto, no hay evidencia suficiente para afirmar que hay una diferencia entre los grupos.

Al finalizar el experimento, se obtuvo un valor de p aproximado a cero en todas las pruebas para cada mejor método pivotado y comparado con los otros métodos, por lo tanto, se concluye que efectivamente existe una diferencia estadísticamente significativa entre los conjuntos de datos.

5.6.4. Entrenamiento del modelo seleccionado utilizando imágenes con mejoras aplicadas

Durante esta etapa, se procede con el entrenamiento utilizando el modelo que demostró el mejor

rendimiento en la fase anterior de evaluación: el YOLOv8 - Nano. Se aplica cada filtro de mejora a los tres conjuntos de imágenes designados para entrenamiento, validación y prueba. Luego, se ejecuta un análisis similar al realizado durante la selección inicial del modelo.

Este proceso incluye la aplicación de cada filtro a los conjuntos de imágenes mencionados, permitiendo una evaluación exhaustiva del impacto de cada filtro en el desempeño del modelo YOLOv8 - Nano. De esta manera, se busca identificar cómo cada modificación influye en la capacidad de detección y localización de tumores en las imágenes. El análisis resultante proporcionará una comprensión más completa sobre cómo los filtros mejoran o afectan el rendimiento del modelo, permitiendo una toma de decisiones más informada en la mejora del sistema de detección de tumores. En la tabla 4 se detallan los resultados.

Tabla 4: Resultados de entrenamientos utilizando el modelo YOLOv8 - Nano y las imágenes mejoradas

Método	Precisión	Sensibilidad	mAP@0,5	mAP@0,5:0,95
CLAHE	0,875	0,811	0,875	0,408
Bilateral	0,827	0,789	0,85	0,388
Suavizado	0,863	0,805	0,852	0,397
BBHE	0,823	0,828	0,861	0,398
Anisotrópico	0,878	0,767	0,859	0,403
Morfológico	0,804	0,755	0,809	0,359
Stretch	0,869	0,784	0,851	0,402

Los resultados obtenidos indican que, en términos generales, el CLAHE destacó con un rendimiento superior en comparación con los otros métodos evaluados. Se destacó al obtener la puntuación más alta en dos de las cuatro métricas consideradas, y en las restantes dos métricas, ocupó el segundo puesto por una diferencia mínima. Específicamente, al analizar la métrica de Precisión, la diferencia fue de apenas un 0,3 %.

A pesar de estos resultados positivos, es importante destacar que ninguno de los métodos logró superar el rendimiento del mismo modelo con las imágenes originales. Este hallazgo resalta la importancia y la complejidad de la tarea, evidenciando que, a pesar de los avances, aún existe margen para mejorar la detección de tumores en las imágenes mediante la aplicación de distintos métodos y filtros.

5.7. Entrenamiento con aumento de datos en línea

El aumento de datos en línea implica la introducción de variaciones aleatorias en la información de entrenamiento en cada época. En este caso, se considera aplicar el filtro con mejor rendimiento del entrenamiento previo, que se llevó a cabo sin el aumento de datos en línea; dicho filtro corresponde al CLAHE.

Se planea continuar con la implementación del aumento de datos en línea mientras se mantiene el uso del modelo que demostró mejor desempeño en las pruebas iniciales, el YOLOv8 - Nano. Este enfoque permite evaluar el efecto específico del filtro pre-

viamente identificado como más efectivo en un escenario donde se aplican variaciones aleatorias adicionales en los datos de entrenamiento durante cada iteración. Esta estrategia busca determinar la contribución del filtro CLAHE en un entorno de entrenamiento que incorpora cambios aleatorios en los datos en cada época. Los resultados se encuentran en la tabla 5

Tabla 5: Resultados de aumento de datos en tiempo real

Método	Precisión	Sensibilidad	mAP@0,5	mAP@0,5:0,95
CLAHE	0,861	0,867	0,903	0,433

En los resultados se puede observar, una mejoría sustancial respecto al entrenamiento sin el aumento de datos en tiempo real, sin embargo, respecto al entrenamiento sin ningún filtro aplicado a las imágenes en el proceso, solo obtuvo un mejor resultado en lo que refiere a la Sensibilidad, donde hay una mejora del 1 %. En el resto de las métricas, hay resultados bastante cercanos, principalmente en las mediciones de los mAP.

6. Conclusión

La culminación de este trabajo de tesis se ha traducido en el desarrollo exitoso de la detección de tumores de cáncer de pulmón en tomografías computarizadas utilizando la arquitectura YOLO y métodos de mejoras en imágenes.

Durante el mismo, se procedió a realizar una prueba de rendimiento de distintos modelos YOLO para luego realizar un análisis de rendimiento y concluir cual era el más apto para la tarea. Luego se implementó distintos filtros de mejora en imágenes, esperando mejorar distintas métricas de las mismas. Con el método de Wilcoxon de a pares se logró concluir que estos filtros afectaban de forma distinta a las métricas de las imágenes.

Finalmente, se realizaron pruebas de aumento de datos utilizando los métodos de mejora de imagen con el modelo que obtuvo el mejor rendimiento en el Experimento 1, con este resultado, se realizó un aumento de datos en línea, para finalmente compararlo con el rendimiento original y analizar los resultados. Las conclusiones se pueden desglosar en los siguientes puntos:

- Evaluación exhaustiva de modelos de detección: Se ha realizado una comparación minuciosa entre los modelos YOLOv7 y YOLOv8, destacando el rendimiento superior del modelo YOLOv8 - Nano. Este modelo logró altos niveles de precisión, sensibilidad y mAP@0.5, demostrando su habilidad para identificar y localizar tumores pulmonares con precisión en imágenes de tomografía computarizada.

- Análisis detallado de técnicas de mejora de imagen: Se evaluaron múltiples métodos de mejora de imagen, resaltando el método CLAHE como el más efectivo en la mayoría de las métricas. Sin embargo, cada método presenta fortalezas en métricas específicas, subrayando la importancia de elegir el método basándose en el aspecto específico que se busca mejorar en las imágenes.
- Análisis mediante el Método de Wilcoxon: Confirmó diferencias estadísticas significativas entre varios métodos, respaldando la importancia de elegir adecuadamente las mejoras aplicadas para mejorar la precisión de los modelos de detección de tumores.
- Comparación y entrenamiento con imágenes mejoradas: Los entrenamientos con los métodos de mejora de imágenes no superaron el rendimiento del modelo con imágenes originales. Esto destaca la complejidad de la tarea y la necesidad de seguir investigando para mejorar la detección de tumores.
- Aplicación de aumento de datos en línea: El uso del método CLAHE durante el aumento de datos en línea demostró mejoras en la sensibilidad del modelo YOLOv8 - Nano. Esta estrategia resalta la importancia de combinar un modelo sólido con datos mejorados para mejorar la detección de tumores.

6.1. Trabajos Futuros

En vista del presente trabajo final de grado, se plantean como futuras investigaciones las siguientes iniciativas que podrían enriquecer el rendimiento actual del proyecto:

- Explorar otras arquitecturas destacadas en la detección de objetos y comparar su desempeño con los resultados obtenidos hasta la fecha.
- Explotar las herramientas adicionales de YOLO para llevar a cabo la clasificación y segmentación de tumores.
- Utilizar técnicas de ciencia de datos para analizar los datos biométricos de los pacientes, con el propósito de mejorar la detección temprana.
- Desarrollar una plataforma orientada a pacientes y profesionales médicos, permitiendo la realización de detecciones e impulsando la telemedicina.
- Adaptar los procedimientos desarrollados para el análisis de imágenes de rayos X, expandiendo así el alcance de la investigación a esta área específica.
- Extender la investigación a conjuntos de datos más diversos y variados, incluyendo datos de diferentes centros médicos o poblaciones, para evaluar la robustez y la generalización del modelo en diferentes contextos.

Referencias

- [1] S. McGuire, "World cancer report 2014," vol. 7, no. 2, pp. 418–419.
- [2] L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent, and A. Jemal, "Global cancer statistics, 2012," vol. 65, no. 2, pp. 87–108.
- [3] J. A. Holmes and R. C. Chen, "Racial disparities in time from diagnosis to treatment for stage i non-small cell lung cancer," vol. 2, no. 1, p. pky007.
- [4] Y. Liu, J. Kim, Y. Balagurunathan, Q. Li, A. Garcia, O. Stringfield, Z. Ye, R. Gillies, M. Schabath, and Y. Ye, "Radiologic features of covid-19 pneumonia in patients with cancer," *Journal of Thoracic Oncology*, vol. 15, pp. 1356–1362, 2020.
- [5] W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, "Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification," *Pattern Recognition*, vol. 61, pp. 663–673, 2017.
- [6] S. Hawkins, J. Korecki, Y. Balagurunathan, Y. Gu, V. Kumar, S. Basu, H. Aerts, A. Dekker, B. Zhao, B. Siegel, A. Devaraj, P. Lee, D. Fried, R. Gillies, L. Schwartz, W. Zhao, S. Napel, M. McNitt-Gray, and D. Goldgof, "Predicting outcomes of non-small cell lung cancer using ct image features," *IEEE Access*, vol. 7, pp. 91 871–91 880, 2019.
- [7] W. Zhu, Y. Huang, L. Zeng, X. Chen, Y. Liu, Z. Qian, N. Du, W. Fan, and X. Xie, "Anatomically-aware multiple organ segmentation for ct using multi-task neural networks," *IEEE Transactions on Medical Imaging*, vol. 38, pp. 2859–2870, 2019.
- [8] J. Zhang, X. Gao, Y. Li, J. Liu, Y. Du, and J. Tian, "Deep learning for lung cancer detection: Tumor delineation and domain adaptation," *Medical Image Analysis*, vol. 58, p. 101547, 2019.
- [9] O. M. de la Salud. (2020) Cáncer. [Online]. Available: <https://www.who.int/es/news-room/fact-sheets/detail/cancer>
- [10] M. Clinic. (2020) Cáncer de pulmón - síntomas y causas. [Online]. Available: <https://www.mayoclinic.org/es/diseases-conditions/lung-cancer/symptoms-causes/syc-20374620>

- [11] R. C. González and R. E. Woods, *Digital Image Processing*, 4th ed. Pearson, 2018. <https://www.softzone.es/programas/juegos/antialiasing-ssa-fxaa-msaa-filtro-anisotropico/>
- [12] J. L. García, "Tomografía computarizada por rayos x: fundamentos y actualidad," *Revista Ingeniería Biomédica*, vol. 2, no. 4, pp. 3–15, 2008.
- [13] IBM. (2020) ¿qué es la inteligencia artificial (ia)? [Online]. Available: <https://www.ibm.com/es-es/topics/artificial-intelligence>
- [14] Wikipedia. (2020) Inteligencia artificial. [Online]. Available: https://es.wikipedia.org/wiki/Inteligencia_artificial
- [15] AWS. (2020) ¿qué es una red neuronal? [Online]. Available: <https://aws.amazon.com/es/what-is/neural-network/>
- [16] IBM. (2020) ¿qué son las redes neuronales convolucionales? [Online]. Available: <https://www.ibm.com/es-es/topics/convolutional-neural-networkse>
- [17] C. de redes neuronales ICHI.PRO. (2020) Capas de redes neuronales - ichi.pro. [Online]. Available: <https://ichi.pro/es/capas-de-redes-neuronales-129624486048658>
- [18] P. Li, S. Wang, T. Li, J. Lu, Y. HuangFu, and D. Wang. A large-scale ct and pet/ct dataset for lung cancer diagnosis (lung-pet-ct-dx) [data set]. the cancer imaging archive. [Online]. Available: <https://doi.org/10.7937/TCIA.2020.NNC2-0461>
- [19] F. Altenberger and C. Lenz, "A non-technical survey on deep convolutional neural network architectures," 03 2018.
- [20] Y. Zhang, Y. Xie, Y. Li, C. Shen, and Y. Xia, "Covid-19 screening on chest x-ray images using deep learning based anomaly detection," vol. 138, pp. 638–643.
- [21] S. Agudo. (2019, 2) Qué es el antialiasing, ssa, fxaa, msaa y el filtro anisótropico. [Online]. Available: <https://www.softzone.es/programas/juegos/antialiasing-ssa-fxaa-msaa-filtro-anisotropico/>
- [22] Wikipedia. (2021) Filtrado anisotrópico. [Online]. Available: https://es.wikipedia.org/wiki/Filtrado_anisotropico
- [23] A. R. R. S. D. Chen, "Minimum mean brightness error bi-histogram equalization in contrast enhancement," vol. 49, pp. 1310–1319.
- [24] M. A. Khan, M. A. U. Khan, , and S. A. Khan, "A comprehensive survey of image enhancement techniques," vol. 5, p. 19, 2019.
- [25] A. G. (2019) What is the entropy of an image and how is it calculated? [Online]. Available: <https://stackoverflow.com/questions/54938747/what-is-the-entropy-of-an-image-and-how-is-it-calculated>
- [26] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," vol. 13, pp. 600–612, 2004.
- [27] A. M. Eskicioglu and P. S. Fisher, "Image quality measures and their performance," vol. 43, pp. 2959–2965, 1995.
- [28] H. H. X. Ma and P. Wu, "A no-reference edge-preservation assessment index for sar image filters under a bayesian framework based on the ratio gradient," vol. 14, p. 856, 2022.
- [29] PineTools. (2021) Ajustar el contraste de la imagen en línea. [Online]. Available: <https://pinetools.com/es/ajustar-contraste-imagen>
- [30] S. S. Kumar and S. S. Kumar. (2013) Measuring the performance of image contrast enhancement technique. [Online]. Available: <https://www.ijert.org/research/measuring-the-performance-of-image-contrast-enhancement-technique-IJERTV2IS120616.pdf>