

Segmentación pulmonar automática en rutina. El problema es la diversidad de datos, mas no un problema de metodología

Angel Yvan Choquehuanca Peraltilla, *Student* Estefany Pilar Huamán Colque, *Student*,
Eduardo Diaz Huayhuas, *Student* Gustavo Raul Manrique Fernandez, *Student*

Abstract—Esta investigación esta basada en el analisis de datos y la importancia por sobre la metodologia de adquisición. La ingente cantidad que se genera de imagenes DICOM nos permite analizar ciertas patologias que a simple vista no pueden ser detectadas o pronosticadas. Sin embargo al tener una ingente cantidad de imagenes, estas pueden no ser entrenadas adecuadamente, originando diversas interpretaciones por cada tipo de Modelos de Segmentacion.

Comparamos cuatro enfoques genéricos de aprendizaje profundo entrenados en varios conjuntos de datos y dos fácilmente disponibles algoritmos de segmentación pulmonar. Realizamos una evaluación de los datos de imágenes de rutina con más de seis enfermedades diferentes, patrones y tres conjuntos de datos publicados.[3]

Usando diferentes enfoques de aprendizaje profundo, los coeficientes de similitud de Dice (DSC) en medios en conjuntos de datos de prueba no variaron más de 0,02. Cuando se entrena en un conjunto de datos de rutina diverso ($n = 36$), un enfoque estándar (U-net) produce un DSC más alto ($0,97 \pm 0,05$) en comparación con la formación en conjuntos de datos públicos como el Lung Tissue Research Consortium ($0,94 \pm 0,13$, $p = 0,024$) o Anatomía 3 ($0,92 \pm 0,15$, $p = 0,001$). Entrenado en datos de rutina ($n = 231$) que cubren múltiples enfermedades, U-net en comparación con los métodos de referencia arrojan un DSC de $0,98 \pm 0,03$ frente a $0,94 \pm 0,12$ ($p = 0,024$).

I. INTRODUCTION

ACTUALMENTE, el aumento y la digitalización de los hospitales, especialmente en el diagnóstico por imagen, y la necesidad de comunicaciones médicas, ha puesto de relieve la necesidad de estandarizar los protocolos de comunicación y los formatos de la información en sanidad.

Uno de los estándares más exitosos hasta la fecha es DICOM (Digital Imaging and Communications in Medicine).

Para la detección de patologías como el COVID-19, se utiliza algoritmos avanzados para la detección de estas anomalías.

En este proyecto final, utilizaremos el algoritmo KDTree que nos permitirá detectar estas anomalías a partir de imágenes DICOM y con ello generar diagnosticos preliminares con el fin de acelerar la atención hospitalaria.[`purdue'rect`, `brown'rect1`, `brown'rect2`, `nasa'report1`].

II. DICOM

DICOM es un protocolo estándar de comunicación entre sistemas de información y a la vez un formato de almacenamiento de imágenes médicas que aparece como solución a los problemas de interoperabilidad entre tipos de dispositivos.



Fig. 1. DICOM

Una imagen médica por sí misma no aporta suficiente información. Para que sea correctamente interpretada es necesario que vaya acompañada de datos del paciente y de la adquisición. Por eso formatos tradicionales como el .jpeg o el .png se quedan cortos.[7]

El trabajo propuesto parte del interés por aplicar conocimientos del área de ingeniería para brindar soluciones a problemas del área de medicina. En particular se trabajará en el desarrollo de una solución que ayude a los profesionales de la salud a subir archivos DICOM (Digital Imaging and Communications in Medicine) o imágenes derivadas de estos, y poder previsualizarlos, elegir cuales son de interés, y almacenarlos en un formato estándar.

DICOM es un estándar reconocido mundialmente para el intercambio de imágenes médicas. Está pensado para el manejo, visualización, almacenamiento, impresión y transmisión de las mismas. Además, incluye la definición de un formato de archivo y de un protocolo de comunicación de red. El protocolo de comunicación es un protocolo de aplicación, el cual hace uso del protocolo TCP/IP para comunicarse entre sistemas. Los archivos DICOM pueden ser compartidos entre dos entidades que sean capaces de recibir imágenes e información del paciente en formato DICOM.

El formato de archivo DICOM tiene la particularidad de tener tanto información de la imagen como información del paciente (metadata). Fue desarrollado con el objetivo de estandarizar el intercambio de imágenes médicas, aún cuando cada fabricante de equipos de imágenes médicas trabaje con su propio enfoque, mientras que formatos como Nifti, Analyze o Minc, apuntan a facilitar y mejorar el análisis de post procesamiento. Como se puede observar, es muy importante tener un estándar que se utilice en todos los hospitales para el mismo o similar examen, ya que esto ayuda a evitar problemas como por ejemplo si el paciente se muda de un hospital a otro.

DICOM también proporciona interconectividad entre diversos sistemas médicos y soporta todas las ramas de la medicina, siendo actualmente un estándar exhaustivo. Además DICOM soporta información de imágenes comprimidas, mediante un mecanismo que permite a documentos con distinto formato ser encapsulados y transformados en un archivo DICOM. Los esquemas de compresión soportados por DICOM incluyen los siguientes: JPEG, Run-Length Encoding (RLE), JPEG-LS, JPEG-2000, MPEG2/MPEG4, y Deflated. Un archivo DICOM encapsulado incluye metadata relacionada al documento nativo y además la metadata necesaria para crear la cápsula DICOM.[6] La innovación del formato de archivo DICOM fue establecer que la información del píxel no puede ser separada de la descripción del procedimiento médico que llevó a la formación de la imagen en sí misma. En otras palabras, una imagen que está separada de su metadata se convierte en una imagen médica sin sentido. Metadata y la información del píxel se juntan en un único archivo, y el encabezado DICOM, además de la información sobre la matriz de la imagen, contiene una descripción muy completa sobre el procedimiento utilizado para generar la imagen, además de contener información relevante sobre el paciente. Por todo esto, el encabezado es dependiente de la modalidad de la imagen y varía en tamaño. Esto trae aparejada una desventaja importante que radica en la necesidad de un software especial para poder abrir los archivos DICOM, ya que los distintos sistemas operativos no reconocen este formato de forma nativa [4]. Los visores DICOM pueden ser software propietario provisto junto al sistema de imágenes médicas, o de terceros. Dentro de estos últimos, existen herramientas de variado desarrollo, pero con la gran desventaja de que son pagas (RadiAnt, MicroDicom, EMV Dicom Viewer, IrfanView, 3DimViewer) mientras otras requieren de conocimientos técnicos para su implementación y puesta en producción (DICOM Web Viewer, Mango). Otra desventaja que tiene el estándar DICOM está relacionada con los datos de entrada de la información. Esto se debe a la posibilidad de incorporar demasiados campos opcionales al encabezado de un archivo por parte de los fabricantes de equipos de imágenes médicas. Esta desventaja condujo a inconsistencias en el completado de estos campos. Algunos atributos dentro del encabezado de imágenes DICOM muchas veces están incompletos, se dejan en blanco o completan con información incorrecta. Todo esto es causado por la flexibilidad en la definición de la construcción de los archivos DICOM. El encabezado de un archivo DICOM consiste en un preámbulo y un prefijo, seguido de una secuencia de elementos de datos. El preámbulo contiene bytes reservados que no son utilizados, mientras que el prefijo contiene el texto ASCII "DICM". Luego del preámbulo y el prefijo viene una secuencia de elementos de datos. Cada uno de ellos consiste en una etiqueta DICOM, un código ASCII de dos caracteres indicando el valor de representación, el tamaño del elemento, y el valor. Los elementos de datos en el archivo están ordenados por el número de etiqueta DICOM. La imagen en sí misma es solamente otro elemento de datos cuyo valor de representación indica si la información corresponde a una sola imagen, varios cuadros de un estudio o un ciclo de video

A. Características del Formato DICOM

El formato DICOM cuenta con objetos IOD (Information Object Definition), formados por la imagen y su información asociada (Son una representación lógica de objetos del mundo real) y DIMSE (DICOM Message Service Element), operaciones que pueden realizarse sobre un objeto. IOD y DICOM forman SOP, la unidad funcional de DICOM.

Un IOD se compone de IEs (Entidades de información) (Hay IE de paciente, de estudio, de serie, de equipo, de imagen...) que a su vez se componen de uno o varios módulos que a su vez se contienen varios atributos. Un atributo se define con nombre, etiqueta, tipo y descripción.

B. Composición del estandar DICOM



Fig. 2. Diagrama del proceso para el análisis de imágenes DICOM

En el estándar DICOM la información se define mediante un modelo que refleja el mundo real. La imagen es el núcleo de información de un fichero DICOM. Cada fichero contiene, además de la imagen, información sobre el paciente (identificación demográfica y de identificación), el estudio en el que se encuadra la toma de la imagen, la serie a la que pertenece la imagen e información sobre la propia imagen.[5]

C. Importancia del estandar DICOM

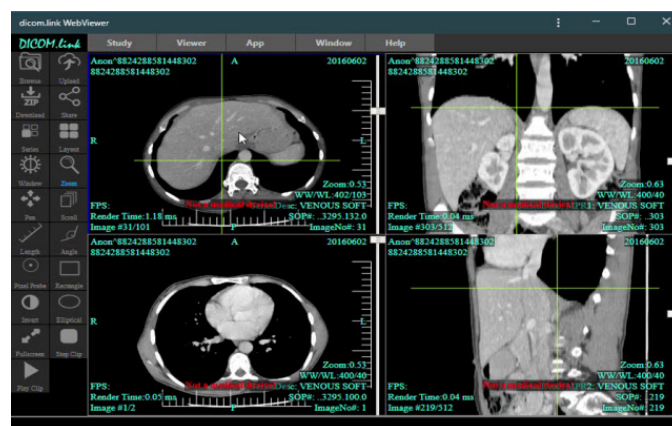


Fig. 3. Diagrama del proceso para el análisis de imágenes DICOM

DICOM permite una identificación unívoca de objetos. Cada fichero DICOM tiene un UID único compuesto por varios números.

Las comunicaciones DICOM se adaptan al estándar OSI para el intercambio de información. La AE (Entidad de Aplicación) se encarga de las comunicaciones de modo que para cada servicio existe un AE cliente y un AE aplicación.

Gracias a sus características y a su nivel de implantación, hoy día DICOM es mundialmente reconocido para el manejo, almacenamiento, impresión y transmisión de imágenes médicas.[2]

D. Importancia de la Escala Hounsfield

La escala de Unidades Hounsfield (símbolo HU del inglés ‘Hounsfield Units’) es el resultado de la transformación de la escala de coeficientes de atenuación lineal de rayos X en una nueva escala en la cual el valor de atenuación del agua destilada en Condiciones Normales de Presión y Temperatura (CNPT) se define como 0 unidades de Hounsfield (HU).[1]

En la tomografía computarizada se hace uso de una escala para determinar la densidad de los tejidos, esta es la escala de unidades Hounsfield(UH), que mediante la atenuación de rayos X se puede terminar rangos entre -1000 a +1000, cuyos valores dependerá del tejido en cuestión siendo -1000 un valor dado para la radiosensibilidad del aire , 0 para el agua destilada y +1000 para tejidos óseos.[4]

TIPO DE TEJIDO	HU
Hueso compacto	≥ 250
Hueso esponjoso	50-300
Sangre coagulada	70-90
Tiroides	60-80
Hígado	50-70
Sangre entera	50-60
Sustancia gris	37-41
Músculo	35-50
Páncreas	30-50
Riñón	20-40
Sustancia blanca	20-34
Plasma	27 ± 2
LCR	5-10
Grasa	-80 hasta -100
Pulmón	-950 hasta -550

III. METODO DE ANALISIS

Entrenamos cuatro modelos genéricos de segmentación semántica desde cero en tres conjuntos de entrenamiento públicos diferentes y un conjunto de entrenamiento recogido de la rutina clínica. Se evaluó en conjuntos de prueba públicos y de rutina datos, incluidos los casos que presentan patologías graves.

A. Rutina de Extracción de Datos

Se tomaron datos clínicos bajo el concepto de ética profesional y anonimato. Con ello se obtuvo datos de imágenes. Para la cantidad de imágenes, se utilizaron imágenes de los pacientes hospitalizados y pacientes ambulatorios que se sometieron a un examen de TC de tórax durante un período de 2,5 años, sin restricción de edad, sexo o indicación. Sin embargo, se aplicó un mínimo de criterio de inclusión con respecto a los parámetros de imagen, como etiqueta DICOM principal y original, número de cortes en una serie mayor o

igual que 100, kernel de convolución aguda, y la descripción de la serie incluía uno de los términos pulmón, tórax o tórax. [8]

B. Entrenando datasets

Para estudiar la diversidad de datos de entrenamiento, se reunió conjuntos de datos con igual número de pacientes con una cantidad específica de pacientes. Luego se hizo una subsección para extraer aleatoriamente datos del conjunto público que sea tipo VISCERAL

Además, seleccionamos cuidadosamente un gran conjunto de datos de entrenamiento representativo de la rutina clínica utilizando tres estrategias de muestreo:

- 1) Muestreo aleatorio de casos (n = 57)
- 2) Muestreo de fenotipos de imagen [28] (n = 71) (la metodología exacta para el fenotipo identificación no estaba en el alcance de este trabajo),
- 3) Selección manual de casos extremos con severa patologías, como fibrosis (n = 28), trauma (n = 20), y otros casos que muestran extensas en vidrio esmerilado opacidad, consolidaciones, patrones fibróticos, tumores y derrames (n = 55)

. En total, seleccionamos 231 casos a partir de datos de rutina para el entrenamiento. Se le asignará el código a los datos como R-231

C. Conjunto de datos de prueba (Test dataset)

Para la prueba, muestreamos aleatoriamente 20 casos de la base de datos de rutinas que no formaban parte del conjunto de entrenamiento y 15 casos con anomalías específicas: atelectasias (n = 2), enfisema (n = 2), fibrosis (n = 4), masa (n = 2), neumotórax (n = 2) y trauma (n = 3).

Se calcularon los resultados en un conjunto de datos combinado compuesto por los conjuntos de prueba individuales (Todos (L), n = 191). Además, reportamos todos los casos de prueba combinados sin los datos LTRC y LCTSC considerados (Todos, n = 62). La razón detrás de esto es que la prueba LTRC conjunto de datos contiene 105 volúmenes y domina el puntaje promedio, y el conjunto de datos LCTSC contiene múltiples casos con tumores y derrames que son no incluido en las máscaras de verdad de tierra. Con esto se evita el descarte de error estadístico.

D. Metodos de segmentación

Nos abstuvimos de desarrollar una metodología especializada, sino que utilizamos arquitecturas genéricas de segmentación semántica y aprendizaje profundo de última generación que eran no propuesto específicamente para la segmentación pulmonar. Se entrenó estos modelos “vanilla”, o sea sin modificaciones y sin entrenamiento previo sobre otros datos. Consideramos las siguientes cuatro segmentaciones semánticas genéricas modelos:

- U-net
- ResU-net
- Dilated Residual NetworkD-22
- Deeplab v3+

E. Modelos de Segmentacion Semanticas

1) *U-net*: Ronneberger (CITA) propusieron la U-Net para la segmentación de estructuras anatómicas en imágenes de microscopía. Desde entonces, se ha utilizado para una amplia gama de tareas de segmentación y varias versiones modificadas han sido estudiadas. Utilizamos la U-Net porque puede trabajar con contenido en lote (batch).

2) *ResU-net*: Se han propuesto conexiones residuales para facilitar el aprendizaje de redes más profundas. El modelo ResU-Net incluye conexiones residuales en cada bajada y bloque de muestreo ascendente como una segunda adaptación al modelo U-net, además de la normalización en lotes.

3) *Dilated Residual Network-D-22*: Yu y Koltun propusieron circunvoluciones dilatadas para segmentación semántica de imágenes y residuos profundos de redes adaptadas con circunvoluciones dilatadas para realizar segmentaciones semánticas sobre imágenes naturales. Se utilizó el modelo Dilated Residual Network-D-22, como propuesto por Yu y Koltun.

4) *DeepLab v3+*: DeepLab v3 combina circunvoluciones dilatadas, multiescala representaciones de imágenes y campos aleatorios condicionales totalmente conectados como un paso de posprocesamiento. DeepLab v3 incluye un módulo decodificador adicional para refinar el segmentación. Aquí, utilizamos el modelo DeepLab v3+ según lo propuesto por Chen. Comparamos los modelos entrenados con dos fácilmente métodos de referencia disponibles: el Holísticamente Progresivo Redes Anidadas (P-HNN) y la Imagen de Tórax Plataforma (CIP). El modelo disponible previa solicitud fue entrenado en casos de la conjunto de datos LTRC público (618 casos) y otros casos con enfermedades pulmonares intersticiales o enfermedades infecciosas (125 casos).

F. Experimentos

Determinamos la influencia de la variabilidad de los datos de entrenamiento (especialmente conjuntos de datos públicos versus rutinarios) en el generalizabilidad a otros conjuntos de datos de prueba públicos y, específicamente, a casos con una variedad de patologías. establecer comparabilidad, limitamos el número de volúmenes y cortes para que coincidan con el conjunto de datos más pequeño de LCTSC, con 36 volúmenes y 3.393 cortes. Durante este experimento, nos consideraron solo cortes que mostraban el pulmón (durante entrenamiento y pruebas) para prevenir un sesgo inducido por el campo de visión. Por ejemplo, imágenes en VISCERAL La anatomía 3 mostraba todo el cuerpo o el tronco, incluyendo el abdomen, mientras que otros conjuntos de datos, como LTRC, LCTSC o VESSEL12, solo contenían imágenes limitado al pecho.

IV. EJECUCIÓN DEL PROYECTO

A. Infraestructura a utilizar

- Google Colab
- Base de Datos de Placas DICOM
- Python 3+
- Biblioteca Lungmask <https://github.com/JoHof/lungmask>
- Biblioteca SimpleITK <https://simpleitk.org/>

B. Diagrama a emplear

C. Repositorio Github de Pruebas

Todos el codigo incluido la base de datos de archivos DICOM se encuentra en el repositorio: Bajo un cuaderno de Colab que se puede ejecutar con los archivos externos para su conversión

D. Uso de la librería Lungmask

1) *Instalación*: Para instalar la librería lungmask se necesita:

- `pip install lungmask` , o
- `pip install git+https://github.com/JoHof/lungmask` , si se necesita la ultima versión.

2) *Ejecución*: Ya instalado previamente, se utiliza la siguiente sintaxis.

- `lungmask INPUT OUTPUT`

La variable INPUT apunta a un archivo, la cual va a ser procesada. Si INPUT apunta a una carpeta, esta carpeta será con contenido de archivos DICOM. Los volúmenes grandes (en terminos de numero de voxels) seran usados para computar usando lungmask. OUTPUT es el nombre de archivo de salida. Todos los formatos ITK son soportados.

Eligiendo un modelo: El algoritmo U-net(R231) es usado por defecto. De todas maneras se puede utilizar los algoritmos que se especifican:

- U-net(R231)
- U-net(LTRCLobes)
- U-net(LTRCLobes R231)
- U-net(R231CovidWeb)

Quedaría así:

- `lungmask INPUT OUTPUT --modelname LTRCLobes`

Para mas ayuda, ejecutar: `lungmask -h`

V. DESARROLLO Y RESULTADOS

Para la pruebas del software descrito en el paper se usó imágenes del tipo DICOM, de un paciente del Centro Médico de Arequipa, las imágenes son resultantes de usar un tomo-grafo del hospital.

El paciente cuenta con 133 imágenes Dicom.

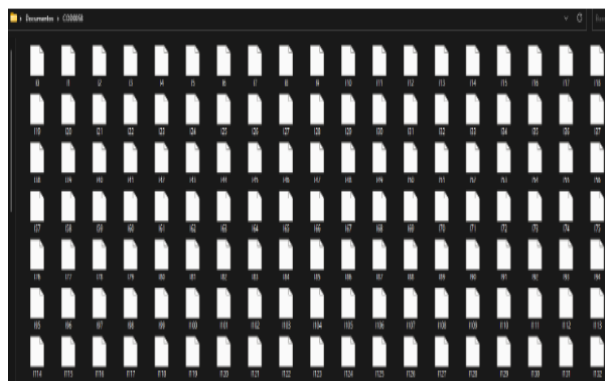


Fig. 4. Dataset de Imagenes Dicom

Por motivos de privacidad se omiten algunos datos sensibles de los pacientes. En el siguiente enlace se encuentra el software usado para anonimizar los datos del paciente. Se utilizó dicho software porque el hospital solo podría proporcionar imágenes sin metadata. <http://www.dclunie.com/pixelmed/software/webstart/DicomCleanerUsage.html>

Las imágenes que obtiene el tomógrafo es de diferentes planos anatómicos entre los cuáles tenemos los axiales, sagitales y coronales.

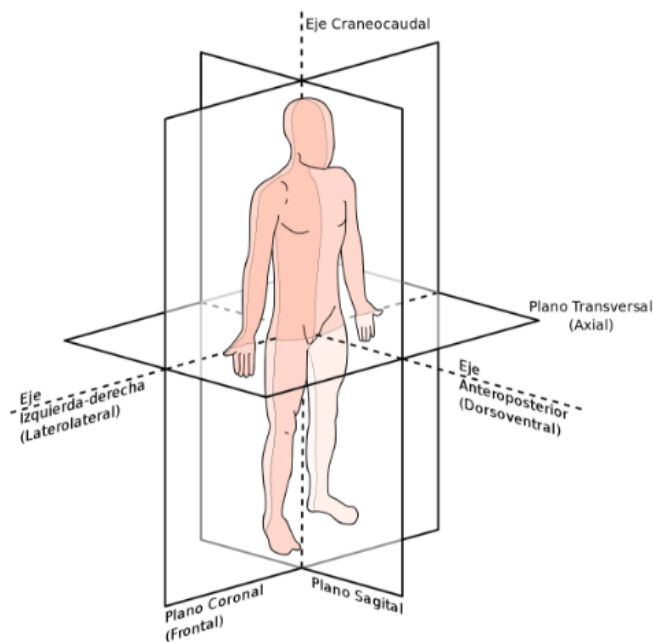


Fig. 5. Planos Anatomicos

El programa elaborado en el artículo solo trabaja con los cortes axiales, dentro de los cuáles solo encontramos 33 cortes. Para detectar los cortes axiales se utilizó una herramienta MITK para visualización de imágenes DICOM.

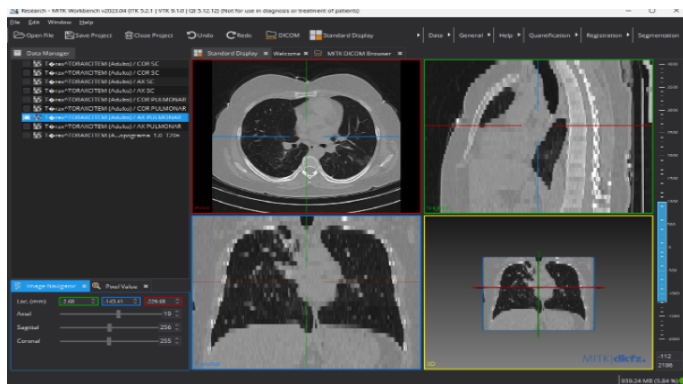


Fig. 6. Herramienta MITK

Luego de detectar las series axiales y la cantidad de cortes de esa serie procedemos a ejecutar el programa del artículo. Para la ejecución del código deberemos colocar las imágenes en la carpeta test/testdata.

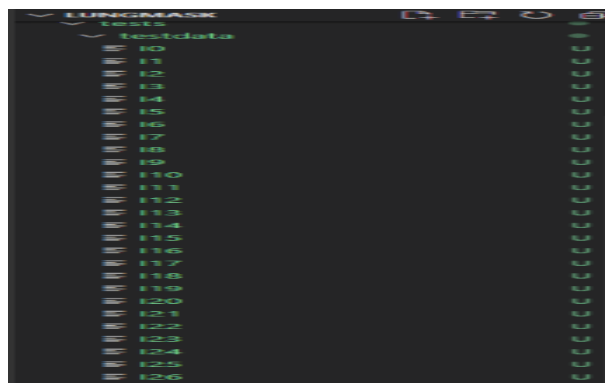


Fig. 7. Carpeta Testdata

Para ejecutarlo instalamos todas las librerías en requirements.txt con el comando:

- `pip install -r requirements.txt`

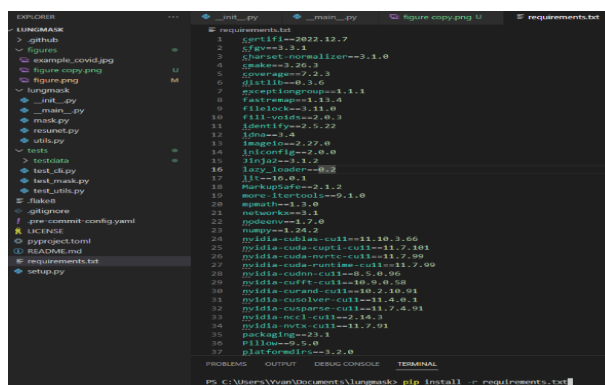


Fig. 8. Dependencias del Proyecto

Para ejecutar el proyecto y se realice la segmentación ejecutaremos el siguiente comando:

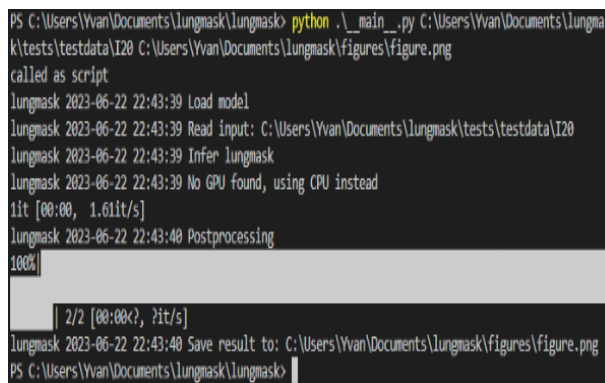


Fig. 9. Segmentacion con Lungmask

En el comando anterior se ejecuta el archivo main.py y se pasa la imagen a segmentar como argumento input y el archivo figure.png como argumento output.

El resultado se almacenará en la carpeta figures, en el archivo figure.png. Se obtendrá un resultado como se muestra a continuación:

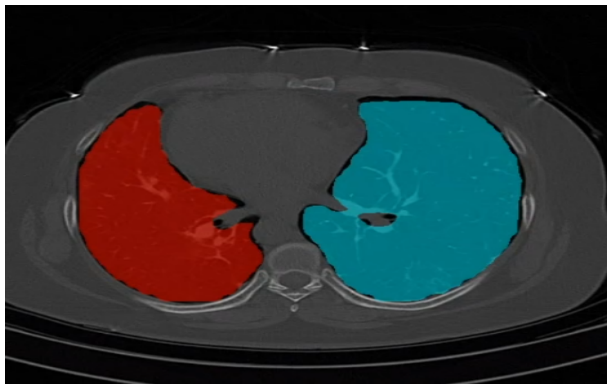


Fig. 10. Archivo Figure.png

Posteriormente podremos realizar la segmentación de todas las imágenes de los cortes axiales del paciente.

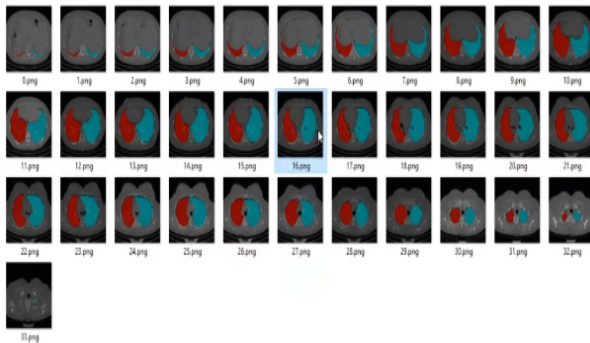


Fig. 11. Resultado de Imagenes Analizadas

- [4] F Shariaty, S Hosseinlou, and V Yu Rud'. "Automatic lung segmentation method in computed tomography scans". In: *Journal of Physics: Conference Series*. Vol. 1236. 1. IOP Publishing. 2019, p. 012028.
- [5] Yasith Sinhabahu Arachchige Don. "COVID-19 Lung Mask Segmentation and Severity Estimation using Deep Learning". PhD thesis. 2022.
- [6] Veronika Stennemark. "Lungmask och löpmagsnematod hos nötkreatur". In: (2011).
- [7] Jef Vandemeulebroucke et al. "Automated segmentation of a motion mask to preserve sliding motion in deformable registration of thoracic CT". In: *Medical physics* 39.2 (2012), pp. 1006–1015.
- [8] Haiyan Wei et al. "An identification methods of under-segmented regions in segmented lung mask". In: *Eleventh International Conference on Graphics and Image Processing (ICGIP 2019)*. Vol. 11373. SPIE. 2020, pp. 60–67.

VI. CONCLUSION

- Se logró aplicar un algoritmo computacional para la detección de patologías en imágenes DICOM
- Se determinó la presencia de patologías en imágenes DICOM mediante el uso de algoritmos computacionales y basados en la escala Hounsfield

REFERENCES

- [1] Peng Gang et al. "Effect of data augmentation and lung mask segmentation for automated chest radiograph interpretation of some lung diseases". In: *Neural Information Processing: 26th International Conference, ICONIP 2019, Sydney, NSW, Australia, December 12–15, 2019, Proceedings, Part IV* 26. Springer. 2019, pp. 333–340.
- [2] Evi Kopelowitz and Guy Engelhard. "Lung nodules detection and segmentation using 3D mask-RCNN". In: *arXiv preprint arXiv:1907.07676* (2019).
- [3] Kokka Paramban Sabeerali et al. "Ai-powered semantic segmentation and fluid volume calculation of lung ct images in covid-19 patients". In: *Emergent Converging Technologies and Biomedical Systems: Select Proceedings of ETBS 2021*. Springer, 2022, pp. 93–101.