



Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

**ANÁLISIS COMPARATIVO DE DATOS DE
TOMOGRAFÍA COMPUTARIZADA DE TÓRAX
DE PACIENTES CON Y SIN COVID-19**

Tesis

presentada al Posgrado de Física Aplicada

como requisito para obtener el grado de

MAESTRO EN CIENCIAS EN FÍSICA APLICADA

por

LUIS GERMÁN PUENTE LAPUERTA*

*luis.puentel@alumno.buap.mx

Asesorado por:

Dr. Cristian Heber Zepeda F.*

Dr. Eduardo Moreno Barbosa†

*hzepeda@fcfm.buap.mx

†emoreno@fcfm.buap.mx

Puebla, 12 de febrero de 2024

Título: Análisis comparativo de datos de tomografía computarizada de tórax de pacientes con y sin covid-19.

Estudiante: Luis Germán Puente Lapuerta.

COMITÉ

Dr. Javier Miguel Hernández López
Presidente

Dr. Benito de Celis Alonso
Secretario

Dr. Lucio Fidel Rebolledo Herrera
Vocal

Dr. Cristian Heber Zepeda Fernández
Asesor

Dr. Eduardo Moreno Barbosa
Co-Asesor



Resumen

Durante el brote de la pandemia de la enfermedad por coronavirus de 2019 (COVID-19), la tomografía computarizada (TC) se convirtió en una forma útil de diagnosticar a los pacientes con esta enfermedad y representa una herramienta importante en la toma de decisiones para los especialistas humanos para tratar de aplicar de una manera oportuna los protocolos médicos adecuados. Pero debido a la privacidad de las imágenes y datos de TC se ha dificultado la investigación y desarrollo de técnicas de inteligencia artificial para la clasificación automática de esta enfermedad y las investigaciones que se han llevado a cabo arrojan resultados erróneos lo que ha impedido su uso clínico. Para el análisis de datos (imágenes) se trabajó con un conjunto de datos (*dataset*) de código abierto en el que se tomaron 133 imágenes diagnosticadas de positivo para COVID-19 y 70 imágenes de negativo para COVID-19 tomando en cuenta el mapeo de **intensidad de píxeles** de las regiones de interés (ROI) de alta densidad que conforman la imagen tomográfica, se utilizó un algoritmo de segmentación para clasificar imágenes de TC en casos de coronavirus y no coronavirus.

Los pacientes con COVID-19 se clasificaron en tres grupos: grupo no crítico (n=47) a los pacientes que se recuperaron de la COVID-19 y fueron dados de alta, grupo crítico (n=45) pacientes ingresados a la unidad de cuidados intensivos (UCI) y grupo de defunción (n=41) a las personas que fallecieron. Se identificó las áreas de mayor densidad en las imágenes de TC, con el fin de ayudar a detectar eficazmente las posibles zonas de anomalías en los pulmones partiendo de la caracterización cuantitativa de la segmentación de los lóbulos pulmonares.



Agradecimientos

A mis padres, quienes siempre han estado conmigo en todo momento. Ellos han sido mi mayor motivación, ya que con su trabajo y su esfuerzo han sabido darnos ese ejemplo de superación y amor a nuestro núcleo familiar. Infinitamente agradecido porque a cada instante me tienen presente en sus oraciones y por me envían bendiciones todos los días.

A ti Mayra, que incluso sabiendo que nos íbamos a alejar 2 años me estimulaste desde el inicio para ir tras uno de mis sueños. Tu eres parte de todo este proceso, gracias por tus palabras de ánimo cuando las cosas no salían bien y por tu regocijo cuando mis propósitos se cumplían. Tu has sido mi sostén y mi compañía durante este trayecto.

A Dios, por darme la salud, los conocimientos, la vida y por permitirme pasar esta linda etapa de mi vida en un país con una enorme diversidad cultural. Me llevaré por siempre un grato recuerdo de sus costumbres y tradiciones.



Índice

Resumen	1
Agradecimientos	3
Índice de tablas	8
Índice de figuras	11
Abreviaturas	12
1. Introducción	13
1.1. Planteamiento del problema	14
1.2. Información sobre la base de datos	14
1.3. Inquietudes y Aclaraciones	15
1.4. Objetivos	16
1.5. Estado del Arte	16
2. Marco Teórico	19
2.1. Historia de los rayos-X	19
2.2. Origen de la Tomografía Computarizada	20
2.3. Radiación electromagnética y producción de rayos-X	21
2.3.1. Energía del fotón	22
2.3.2. Tubo de rayos-X	24
2.4. Interacción de rayos-X con la materia	25
2.5. Píxel y Vóxel	26
2.6. Coeficiente de Atenuación lineal (μ)	27
2.7. Unidades Hounsfield (UH)	28
2.7.1. Fórmula	29
2.8. Contraste de una imagen	29
2.9. Tomografía Lineal	30
2.10. Tomografía Computarizada (TC)	30



2.11. Imágenes Digitales	32
2.11.1. Intensidad	33
2.11.2. Muestreo y Cuantificación de Imágenes	33
2.11.3. Representación digital de imágenes	35
2.11.4. Resolución	37
2.11.5. Profundidad de color	38
2.11.6. Píxeles vecinos	39
2.12. Digitalización y Análisis de Imágenes	40
3. Base de datos	43
3.0.1. Imágenes con opacidades circulares	43
3.0.2. Imágenes con regiones faltantes	45
3.0.3. Imágenes con regiones negras	47
3.0.4. Regiones Opacas	48
3.1. Imágenes No COVID-19	49
4. METODOLOGÍA	51
4.1. Procesamiento de Imágenes	51
4.2. Generación de la matriz	51
4.3. Escala de grises	52
4.4. Paleta de color	53
4.5. Imágenes de TC de tórax, mapas de calor e histogramas	54
4.5.1. Imágenes de TC de tórax para el diagnóstico de COVID-19	54
4.5.2. Mapas de Calor	56
4.5.3. Intensidades de las imágenes como variables aleatorias	57
4.5.4. Histogramas de Intensidad	58
4.5.5. Propiedades estadísticas de un histograma: Media y desviación estándar	60
4.6. Segmentación de imágenes	61
4.7. Histogramas de Ruido	63
4.8. Región de Interés (ROI)	65



5. Resultados	67
5.1. Resultados Imágenes pacientes COVID-19	67
5.1.1. Resultado General para pacientes COVID-19	69
5.1.2. Intervalos de Confianza	71
5.2. Resultados Imágenes pacientes No COVID-19	74
5.2.1. Resultado General para pacientes No COVID-19	76
5.3. Comparación entre los tres grupos de COVID-19 vs. los tres grupos de NO COVID-19	76
6. Análisis de Resultados	77
Comentarios finales	78
A. Tabla 11	83
B. Tabla 12	84
C. Código: ImageToHisto()	85
D. Código: gauss medias()	89



Índice de tablas

1.	Tres grupos en función de los resultados clínicos de los pacientes	17
2.	Rango aproximado de Unidades Hounsfield para diferentes tejidos	29
3.	Distintos valores de bits para representar el color píxel en una imagen.	39
4.	Resumen del algoritmo para obtener las máscaras de las imágenes TC	63
5.	Características básicas de los 133 pacientes con COVID-19 estudiados.	67
6.	Valores de intensidad de pacientes diagnosticados con COVID-19	70
7.	Valores de intensidad de pacientes sin COVID-19	71
8.	Nivel de confianza de los grupos analizados	72
9.	Valores de intensidad de los ruidos y de la región acotada de pacientes NO COVID-19	74
10.	Hallazgos radiológicos característicos de pacientes con COVID-19	78
11.	Ruido oscuro de pacientes positivo COVID-19	83
12.	Ruido claro de pacientes positivo COVID-19	84

Índice de figuras

1.	La primera fotografía de rayos-X fue una imagen de la mano izquierda (incluyendo un anillo) de Bertha esposa de Wilhelm Röntgen. Würzburg, 22 de diciembre de 1895. <i>Fuente:</i> American Institute of Physics	20
2.	Espectro electromagnético. De acuerdo a la energía por fotón los rayos gamma son los más energéticos y las ondas de radio son las de menor energía.	22
3.	Esquema convencional de un tubo de rayos-X	24
4.	Interacción de rayos-X con un paciente	26
5.	Unidades Hounsfield (UH), cuyo rango va desde -1000 a 1000	28
6.	Valores de UH del pulmón, grasa (tejido adiposo) y hueso de una TC de tórax de paciente con COVID-19	30
7.	Adquisición de imágenes usando una fuente giratoria de rayos-X	31
8.	Muestreo y cuantificación de una imagen	34



9.	A la izquierda, imagen continua proyectada sobre una matriz de sensores. A la derecha, la imagen luego del muestreo y cuantificación.	35
10.	Convención de coordenadas usadas para representar una imagen digital.	36
11.	Imágenes de TC de tórax adquiridas durante la pandemia por COVID-19	37
12.	Rango de valores en escala de grises	39
13.	Vecinos de un píxel p	40
14.	Paciente Masculino con opacidades circulares 1	43
15.	Paciente Masculino con opacidad circular 2	44
16.	Paciente Masculino con opacidad circular 3	44
17.	Paciente Masculino con opacidad circular 4	45
18.	Paciente Masculino con áreas blancas	45
19.	TC de paciente Masculino con regiones faltantes	46
20.	Paciente Masculino 20	46
21.	Paciente Masculino 21	47
22.	TC paciente masculino con regiones negras	47
23.	TC paciente Masculino 23	48
24.	TC paciente Masculino 24	48
25.	TC paciente Masculino 25	48
26.	Paciente No-Covid 26	49
27.	Paciente No-Covid 27	49
28.	Paciente No-Covid 28	50
29.	Paciente No-Covid 29	50
30.	Paciente No-Covid 30	50
31.	Escala de grises, $z \in [0, 255]$	53
32.	Saturación del color naranja, entre más nos acerquemos a 1 el color será más puro.	53
33.	Paleta de colores usada en esta investigación con un rango de $[0,1]$	54
34.	Cuatro densidades básicas de una TC	55
35.	Transformación de una TC original a un mapa de calor	57
36.	Imagen de TC oscura y su correspondiente histograma. El eje horizontal del histograma son valores de z_k y el vertical son valores de $p(z_k)$	59



37.	Imagen de TC clara y su correspondiente histograma. El eje horizontal del histograma son valores de z_k y el vertical son valores de $p(z_k)$	60
38.	Imagen original con su respectiva redigitalización e histograma.	61
39.	Imagen de TC con el resultado de la segmentación	62
40.	Función máscara. Para este ejemplo el umbral está en el intervalo $[0.45, 0.56]$.	63
41.	Histogramas de Ruido	64
42.	a) Imagen de TC, b) Histograma de la ROI	65
43.	Comparación de las ROI (de color verde) con el ruido oscuro y claro usando ROOT e ImagenJ	66
44.	Histograma de la región de interés del grupo no crítico de una muestra de 47 pacientes.	68
45.	Histograma de la región de interés del grupo crítico de una muestra de 45 pacientes.	69
47.	Histograma de la región de interés de los tres grupos estudiados de una muestra de 133 pacientes.	69
46.	Histograma de la región de interés del grupo de defunción de una muestra de 41 pacientes.	70
48.	Intervalo de confianza	71
49.	Esquema de intervalos de confianza de los grupos de COVID-19 para un nivel de confianza del 95 %	72
50.	Se aprecia los intervalos de confianza de los grupos de COVID-19 para un nivel de confianza del 90 %	73
51.	Diagrama de caja y bigotes donde se observa la dispersión de los datos.	73
52.	Histograma que compara los resultados del grupo crítico con el grupo de corte 1	74
53.	Histograma que compara los resultados del grupo de muerte con el grupo de corte 2	75
54.	Histograma que compara los resultados del grupo no crítico con el grupo de corte 3	75
55.	Histograma de la región de interés de los tres grupos estudiados de pacientes diagnosticados de negativo para COVID-19.	76



56. Histograma de la región de interés de los seis grupos analizados tanto de pa-
cientes con COVID-19 como de pacientes sanos.. 76



Abreviaturas

- **COVID-19:** Enfermedad por coronavirus de 2019 (por sus siglas en inglés de *Coronavirus Disease of 2019*).
- **SARS-CoV-2:** Coronavirus del Síndrome Respiratorio Agudo Severo 2 (por sus siglas en inglés de *Severe Acute Respiratory Syndrome Coronavirus 2*).
- **RT-PCR:** Reacción en Cadena de la Polimerasa con Transcripción Inversa (por sus siglas en inglés de *Reverse Transcription Polymerase Chain Reaction*).
- **TC:** Tomografía Computarizada.
- **TAC:** Tomografía Axial Computarizada.
- **RX:** Rayos X.
- **UH:** Unidades Hounsfield.
- **ROI:** Región de interés (por sus siglas en inglés de *Region of Interest*).
- **GGO:** Opacidad de vidrio esmerilado (por sus siglas en inglés de *Ground-Glass Opacity*).
- **SDRA:** Síndrome de dificultad Respiratoria Aguda.



Capítulo 1

1. Introducción

Hasta enero de 2024 la enfermedad por corona virus 2019 (COVID-19) ha causado la muerte de 6.9 millones de personas alrededor del mundo. Para el año 2020 el método de diagnóstico estándar fue la prueba de reacción en cadena de la polimerasa con transcripción inversa (RT-PCR). Una de las variantes más usadas en la detección temprana y manejo oportuno de pacientes positivos a COVID-19 se fundamenta en el análisis de imágenes, específicamente los especialistas se basan en estudios radiológicos como lo es la tomografía axial computarizada (TAC), también conocida simplemente como tomografía computarizada (TC) de tórax. Tanto la TC como las diferentes técnicas de adquisición de imágenes médicas producen imágenes de áreas del interior del cuerpo y no siempre es una tarea fácil la interpretación de dichas imágenes incluso para los radiólogos expertos, por esta razón, el procesado de imágenes es una solución óptima para el diagnóstico e interpretación de las imágenes médicas.

A pesar de que la TC es muy útil en el diagnóstico de esta enfermedad no ha habido una investigación y desarrollo importante de modelos basados en inteligencia artificial para que ayuden a los radiólogos y médicos en la lectura y apreciación de las imágenes de TC, esto se da por la privacidad de información de los pacientes. Para abordar este problema trabajaremos con una base de datos compuesta de 133 imágenes que fueron previamente diagnosticadas de positivo para COVID-19 y 70 imágenes diagnosticadas para negativo de COVID-19 las cuales incluyen información del paciente como: edad, genero, historia médica, reporte radiológico, entre otras, para construir un modelo de diagnóstico (binario) capaz de extraer los patrones característicos de las lesiones pulmonares de dicha enfermedad y diferenciarlas de pacientes que no padezcan de esta enfermedad viral. Para lograr identificar las áreas afectadas nos apoyamos en un algoritmo escrito en C++.

El proceso comprende la técnica de *análisis de imágenes* llamada **segmentación de imágenes**, donde la entrada es una imagen de TC y la salida también puede ser una imagen o una gráfica de algún atributo extraído de la imagen original. Los algoritmos extraen de las imágenes



de TC códigos o números que están relacionados con la intensidad de los píxeles, que son con los que se trabajarán a lo largo de esta investigación. Por medio de esta extracción de información basada en el valor del nivel de gris de los píxeles es posible dividir a las imágenes hasta un nivel de subdivisión en la que se puedan aislar las regiones de interés (ROI, por sus siglas en inglés de *region of interest*) lo que hace fiable el análisis.

Segmentar una imagen médica muchas veces no es una tarea fácil debido a las semejanzas entre el fondo y las ROI's por lo que es necesario un procesamiento de imágenes adecuado. Además, bajo ciertas consideraciones se elimina el mayor **ruido** posible de las imágenes. Se espera obtener una herramienta de clasificación sofisticada para el diagnóstico de COVID-19 que cumpla con las métricas más extendidas para evaluar el rendimiento del modelo de clasificación con una precisión y exactitud óptimas.

1.1. Planteamiento del problema

En una imagen de TC, los distintos tejidos del cuerpo absorben la radiación en diferentes *intensidades*. Los diferentes valores de intensidad de los píxeles que conforman una imagen de TC de tórax se muestran a través de una escala de grises. Mediante el análisis de estos valores se irá caracterizando las ROI de dichas tomografías tanto de pacientes con diagnóstico de COVID-19 como de pacientes sanos.

Las TC que se estudian son monocromáticas (un solo color), están en la escala de grises, es decir, contiene tonos de gris (una escala continua de negro a blanco). Es así que las variaciones de densidad de tejido están representadas por los niveles de brillo de los píxeles que la conforman, es decir, por el valor numérico correspondiente, que en imágenes monocromáticas van desde 0 = negro hasta 1 = blanco.

1.2. Información sobre la base de datos

El conjunto de datos que se utilizó en la presente tesis es de código abierto y se le puede descargar en el siguiente enlace: <https://github.com/UCSD-AI4H/COVID-CT>, para las TC de pacientes con COVID-19 se seleccionaron manualmente las imágenes que contienen información clínica relevante e historial clínico de los pacientes. Se descartaron imágenes ya



sea por no contener información del paciente o por su pobre calidad. Yang, He Zhao (2020) creadores de esta base de datos recolectaron las imágenes TC de prepublicaciones (*preprints*) sobre COVID-19 como los es medRxiv y bioRxiv. La utilidad de este conjunto de datos ha sido confirmada por un radiólogo senior en el Hospital Tongji, Wuhan, China, que realizó el diagnóstico y tratamiento de una gran cantidad de pacientes con COVID-19 durante el brote de esta enfermedad entre enero y abril.

El formato de archivo de las imágenes es PNG (*Portable Network Graphics*) que muestra imágenes digitales de alta calidad debido a que ofrece una paleta de colores más amplia y brillante (ver sección 4.4), este formato admite una profundidad de color de hasta 24 bits. Permite la compresión sin pérdida de datos favoreciendo así el procesamiento en diferentes situaciones. Este tipo de formato libre de patentes facilita el intercambio, ya sea por medio de dispositivos (memorias) o redes, reduciendo tiempos de transmisión [1].

1.3. Inquietudes y Aclaraciones

Cuando se trabaja con imágenes de TC que fueron extraídas de la web surgen dos inquietudes: en primer lugar, cuando las imágenes de TC originales se ponen en artículos científicos, la calidad de estas imágenes se degrada, lo que puede hacer que las decisiones diagnósticas sean menos precisas. La degradación de la calidad puede incluir que se disminuya el número de bits por píxel y en consecuencia la resolución de las imágenes se reduzca; en segundo lugar, la tomografía computarizada original contiene una secuencia de cortes de TC (CT slides), pero cuando se pone en artículos, solo se seleccionan unos pocos cortes clave, lo que también puede tener un impacto negativo en el diagnóstico.

Pero según radiólogos experimentados, las cuestiones planteadas en estas preocupaciones no afectan significativamente la precisión de la toma de decisiones de diagnóstico. En primer lugar, los radiólogos con experiencia pueden realizar un diagnóstico preciso a partir de imágenes de TC de baja calidad, por ejemplo, dada una foto tomada con un teléfono inteligente de la imagen de TC original. Del mismo modo, la diferencia de calidad entre las imágenes de TC en los artículos y las imágenes de TC originales no perjudicará en gran medida la precisión del diagnóstico. En segundo lugar, si bien es preferible leer una secuencia de cortes de TC,



a menudo un solo corte de TC contiene suficiente información clínica para tomar decisiones precisas [2].

1.4. Objetivos

Caracterizar el tejido sano y tejido con lesiones presentes en los pulmones, estas zonas serán denominadas regiones de interés (ROI), a fin de distinguirlo del ruido (regiones que no son objeto de nuestro análisis, bordes o regiones ajenas a los pulmones). Para este propósito, se utilizará un algoritmo escrito en C++.

Realizar un análisis de datos cuantitativo de las intensidades (densidades) de TC de tórax, para disminuir la subjetividad en el diagnóstico de pacientes con COVID-19.

1.5. Estado del Arte

A partir de que se detectó por primera vez la enfermedad por coronavirus 2019 (COVID-19) en Wuhan, China, considerado el primer epicentro de la pandemia y el hecho importante de que el virus podría transmitirse entre humanos, cientos de científicos, médicos y expertos epidemiólogos alrededor del mundo comenzaron las investigaciones y monitoreo en torno al COVID-19, principalmente en entender lo más pronto posible los efectos de las variantes emergentes, diagnósticos, tratamientos y las eventuales vacunas. El 30 de enero de 2020 la Organización Mundial de la Salud (OMS), declaró la epidemia de COVID-19 como una “emergencia de salud pública de preocupación internacional” y más adelante el 11 de marzo de 2020 el Director General de la OMS anunció a la enfermedad del coronavirus como pandemia mundial.

Una de las primeras regiones en las que el virus se propagó fuera de China fue en los países de Oriente Medio, específicamente en Irán, donde el 19 de febrero de 2020 se confirmó dos casos de muerte causada por el coronavirus del síndrome respiratorio agudo severo 2 (SARS-CoV-2), por sus siglas en inglés de *Severe Acute Respiratory Syndrome Coronavirus 2*. Los trabajos iniciales que se llevaron a cabo fue investigar las características y hallazgos radiológicos principales de imágenes tomográficas de COVID-19 [3]. Se estudiaron los datos e historias clínicas de 363 imágenes de TC de casos confirmados de COVID-19 en el Hospital Baqiyata-



llah situado en Teherán, capital de Irán. Los autores clasificaron a estos pacientes en grupos de acuerdo a la gravedad de su cuadro clínico (ver Tabla 1).

Grupos	Número de pacientes (n)
Pacientes que se han recuperado de COVID-19	194
Pacientes ingresados a unidad de cuidados intensivos	65
Pacientes que han fallecido	104

Tabla 1: Tres grupos en función de los resultados clínicos de los pacientes

Para el análisis estadístico de los datos continuos de los tres grupos se usó la prueba de análisis de varianza unidireccional (ANOVA) un método que consiste en comparar las varianzas entre las medias y para la distribución de los datos categóricos se utilizó la prueba del chi-cuadrado. Para el estudio de los factores de riesgo de los tres grupos se empleó un modelo de regresión logística multinomial. Según este análisis se asociaron a una alta probabilidad de riesgo de muerte la combinación de opacidades en vidrio esmerilado (GGO, por sus siglas en inglés de *Ground-Glass Opacities*) con consolidación.

Otra investigación importante que se ha llevado a cabo fue realizado por Pezzotti con el trabajo *Interpretación de la radiografía de tórax: algo más que blanco y negro* en donde se realiza una diferenciación de los objetos en función de sus densidades relativas [4]. Este autor realiza un estudio minucioso de las alteraciones frecuentes de los tres lóbulos pulmonares del pulmón derecho y de los dos lóbulos del pulmón izquierdo.

Un patrón muy frecuente de enfermedad pulmonar es el edema pulmonar que es la acumulación de líquido en los alvéolos y en una Rx de tórax se observa como patrón en *alas de mariposa*. Este patrón hace que el área central de los pulmones tenga un color blanco [4]. A medida que este patrón muestre progresión se lo apreciará en la Rx de tórax como patrón en *vidrio deslustrado* (uno de los hallazgos más comunes que presentan los pacientes con COVID-19).



Capítulo 2

2. Marco Teórico

2.1. Historia de los rayos-X

Para iniciar con el estudio del procesamiento de imágenes es conveniente comprender como se originan. La principal fuente de energía actualmente es el espectro de energía electromagnética (EM), y en el caso específico de las imágenes de TC su fuente son los **rayos-X** que se encuentran entre las fuentes de radiación EM más antiguas para obtener imágenes, por esta razón es conveniente conocer un poco de su historia.

A finales del siglo XIX en la Universidad de Würzburg surgiría uno de los grandes acontecimientos de la ciencia. En 1895, el científico alemán Wilhelm Conrad Röntgen que se encontraba estudiando las propiedades y fenómenos de los rayos catódicos (haz de electrones que viajan en un tubo al vacío desde un extremo con carga negativa hacia otro extremo con carga positiva) descubrió al replicar los experimentos previos de Hertz y Lenard un tipo de radiación desconocida para aquella época. Encontró que cuando electrones de alta energía chocan contra un objetivo metálico (como por ejemplo el wolframio) estos colisionan con los electrones del metal lo que produce una radiación que penetra en objetos que son opacos a luz.

En una habitación completamente a oscuras, Röntgen hizo pasar corriente de alta tensión a través de un tubo al vacío el cual estaba forrado de cartón negro y observó una extraña luminiscencia sobre un papel recubierto con platinocianuro de bario que se encontraba cerca del tubo. Alejó el aparato a dos metros de distancia y todavía el papel se iluminaba intensamente volviéndose fluorescente con cada descarga eléctrica. Se percató que el delgado cartón era penetrable para estos misteriosos rayos que iluminaban el papel, probó con otros objetos (madera, aluminio, etc.) y observó que también dejaban pasar estos rayos. Sin embargo, no atravesaban el plomo (Pb). Para demostrar que los rayos-X atravesaban diferentes materiales utilizó placas fotográficas. Un día le pide a su esposa (Bertha) colocar su mano entre el aparato y la placa y observó que además de la sombra de la mano se podía apreciar la “sombra” de los huesos de los dedos y los metacarpianos. Röntgen imprimió esta imagen que quedó registrada en la



Figura 1: La primera fotografía de rayos-X fue una imagen de la mano izquierda (incluyendo un anillo) de Bertha esposa de Wilhelm Röntgen. Würzburg, 22 de diciembre de 1895. *Fuente:* American Institute of Physics

placa fotográfica (Figura 1) la que se considera la primera radiografía conocida de una parte del cuerpo humano de la historia.

Más tarde a esta desconocida radiación, Röntgen la denominó “**rayos X**”. Por este excepcional trabajo que es uno de los mayores hallazgos de la ciencia moderna el físico alemán recibió en 1901 el premio Nobel de Física. El descubrimiento de los rayos-X marcó un antes y un después en la ciencia y particularmente en medicina ya que dio paso al inicio de un vasto campo de investigación en técnicas de diagnóstico por imagen y en el surgimiento de la radiología.

2.2. Origen de la Tomografía Computarizada

La tomografía computarizada (TC) en sí se planteó a inicios de la década de 1960. Pero fue en el periodo entre 1970 a 1972 que se presentaron los mayores descubrimientos y avances en TC de forma independiente por parte del autodidacta y técnico eléctrico Godfrey Hounsfield y del físico Allan McLeod Cormack, y que al igual que con el descubrimiento de los rayos-X por parte de Röntgen este invento pasó a revolucionar muchos campos de la ciencia principal-

mente la Medicina y la Física y considerado como uno de los hallazgos más importantes del siglo XX. Hounsfield y el profesor Allan Cormack compartieron en 1979 el Premio Nobel de Medicina por sus aportaciones en la creación del primer sistema de tomografía computarizada para diagnóstico clínico.

Inicialmente fue descrita como **tomografía axial computarizada (TAC)**, en inglés como *computed axial tomography (CAT)* o simplemente conocida como *CAT scanning*. La palabra **axial** fue incluida para diferenciarla de las técnicas de tomografía convencional y el término **computarizada** hace referencia al rol clave que desempeña la computadora en el desarrollo de esta tecnología. Y es que con la aparición de la computadora digital se pudo poner en práctica toda la teoría sobre los métodos de reconstrucción de imágenes que habían sido desarrolladas durante décadas. Y con los enormes avances de la informática a lo largo de los años esta técnica se ha ido transformando desde imágenes transaxiales básicas a una sofisticada representación en tres dimensiones (3D) de diferentes partes del cuerpo.

Desde su aparición hasta la actualidad la TC se ha convertido en una de las principales herramientas para el diagnóstico y exploración de muchas enfermedades, como por ejemplo, en el tratamiento del cáncer, tumores y derrames cerebrales, enfermedades de los pulmones, corazón, etc. Durante el brote de COVID-19 a nivel mundial, la tomografía computarizada (TC) se aprovechó de forma útil en el diagnóstico de pacientes con esta enfermedad y representa una herramienta importante en la toma de decisiones para los especialistas humanos para tratar de aplicar de una manera oportuna los protocolos médicos adecuados.

2.3. Radiación electromagnética y producción de rayos-X

Como lo hemos visto las imágenes de TC tienen su fuente en los rayos-X que son una forma de radiación electromagnética (EM). La principal fuente de energía que se utiliza actualmente es el espectro de energía electromagnética, otra fuente de energía puede ser, por ejemplo, el ultrasonido. La EM puede definirse como la energía (en forma de ondas o partículas elementales) que se propaga ya sea en el vacío o a través de un medio. Todas las formas de radiación electromagnética viajan en el vacío con la misma velocidad de la luz ($c = 3 \times 10^8 \text{ m/s}$) y se agrupan de acuerdo a sus longitudes de onda (λ), frecuencia (ν) o la energía de uno de sus fotones que la

componen (ver Figura 2). En otras palabras, se puede considerar a la radiación electromagnética como campos eléctricos y magnéticos que varían en forma de ondas sinusoidales, dada por la siguiente ecuación

$$c = \lambda\nu \tag{1}$$

donde observamos que ν es inversamente proporcional a λ . También, es de interés la forma en la cual son producidas y como cambian sus propiedades especiales cuando cambian su longitud de onda. Por ejemplo, los rayos-X (emitidos por tubos de rayos-X) se producen en las capas externas del átomo y los rayos gamma, que se producen en el núcleo (radioactivo) del átomo tienen básicamente las mismas propiedades y se diferencian únicamente en su origen y en su frecuencia.

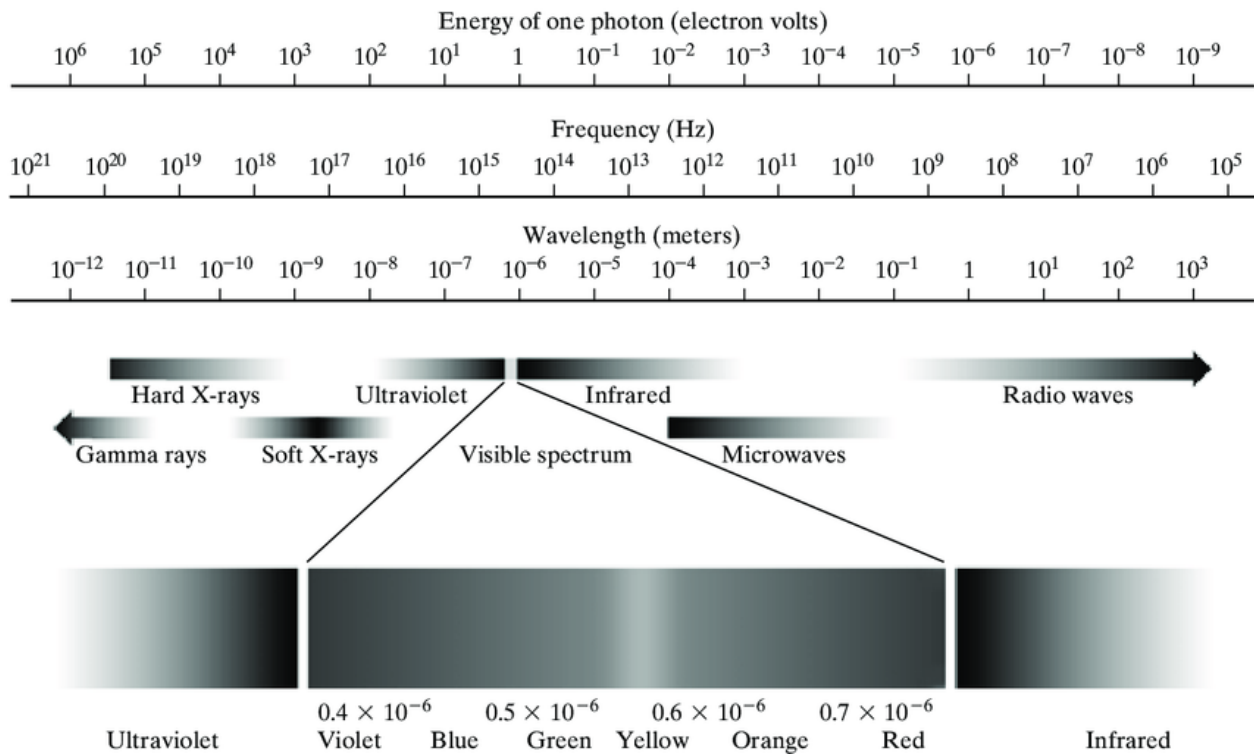


Figura 2: Espectro electromagnético. De acuerdo a la energía por fotón los rayos gamma son los más energéticos y las ondas de radio son las de menor energía.

2.3.1. Energía del fotón

La radiación electromagnética presenta un fenómeno cuántico o también llamado dualidad onda-partícula. Se puede considerar que tiene propiedades de onda pero también se la representa

como paquetes o cuantos de energía llamados **fotones** que son partículas que no tienen masa ni carga eléctrica [5]. Un átomo puede ser excitado de un estado de energía inferior a uno superior de varias maneras, siendo los dos métodos más comunes el impacto de electrones y la *absorción* de la radiación electromagnética. Cuando los átomos excitados vuelven a los estados de menor energía emiten una cantidad de energía que es igual a la diferencia de la energía inicial del estado E_i y el estado final E_f del átomo. Cada átomo libera su exceso de energía en un solo fotón [6]. La frecuencia ν de la radiación emitida es

$$\nu = \frac{E_i - E_f}{h}, \quad (2)$$

donde h es una constante de proporcionalidad llamada *constante de Planck*. Entonces, cuando la radiación interacciona con la materia se comporta como partícula con una energía $E = E_i - E_f$, reemplazando este valor en la anterior expresión obtenemos la ecuación de la energía de todos los tipos de radiación del espectro electromagnético, observamos que la energía es directamente proporcional a la frecuencia

$$E = h\nu. \quad (3)$$

Tanto los rayos-X y los rayos gamma (al igual que todas las formas de radiación electromagnética) están conformados de fotones. Por lo tanto, los rayos-X son paquetes o cuantos de alta energía que se desplazan en ondas con una frecuencia característica. La radiación electromagnética de alta energía, principalmente en las bandas de los rayos-X y de los rayos gamma puede ser perjudicial para ciertos organismos vivos. Una de las propiedades fundamentales de estos rayos es su capacidad de atravesar objetos o materia orgánica y ser registrada en una placa fotográfica, por esta razón que una de sus principales aplicaciones es el diagnóstico médico. Debido a la diferencia de la densidad de los tejidos llegan en mayor o menor proporción a la placa, produciendo una imagen en tonos negros, grises y blancos (escala de grises).

Explicaremos más adelante como los rayos-X son producidos (en el tubo de rayos-X) cuando electrones de alta energía son acelerados por un voltaje de miles de voltios y dirigidos a un blanco (generalmente un material de tungsteno) donde son frenados repentinamente.

2.3.2. Tubo de rayos-X

El tubo de rayos-X es el lugar donde se originan estos rayos de frecuencia comprendida entre 10^{16} y 10^{20} Hz por medio de un procedimiento que consiste en acelerar electrones para luego frenarlos bruscamente obteniéndose así los fotones que establecen la **radiación ionizante** empleada en la adquisición de imágenes de uso clínico. Se dice que es un tipo radiación ionizante ya que al interactuar con la materia es capaz de eliminar un electrón de un átomo o molécula y provocar su ionización.

La estructura fundamental del mencionado tubo representado en la Figura 3 consta de dos electrodos sellados en una envoltura de vidrio al vacío para favorecer que el movimiento de los electrones sea lo más rectilíneo posible [5]. El primer electrodo incorpora una fina bombilla o filamento metálico (cátodo) generalmente de Wolframio (W) por sus buenas propiedades de emisión termoiónica, efecto que ocurre cuando al volverse incandescente emite una nube de electrones a su alrededor (poner cita libro). Este filamento es calentado al pasar una corriente eléctrica hasta una temperatura en la que esté al rojo vivo, a estas altas temperaturas una fracción de electrones son liberados y atraídos por el segundo electrodo denominado ánodo (cargado positivamente) o también conocido como blanco (generalmente de W) acoplado a una pieza de cobre que actúa como disipador de calor.

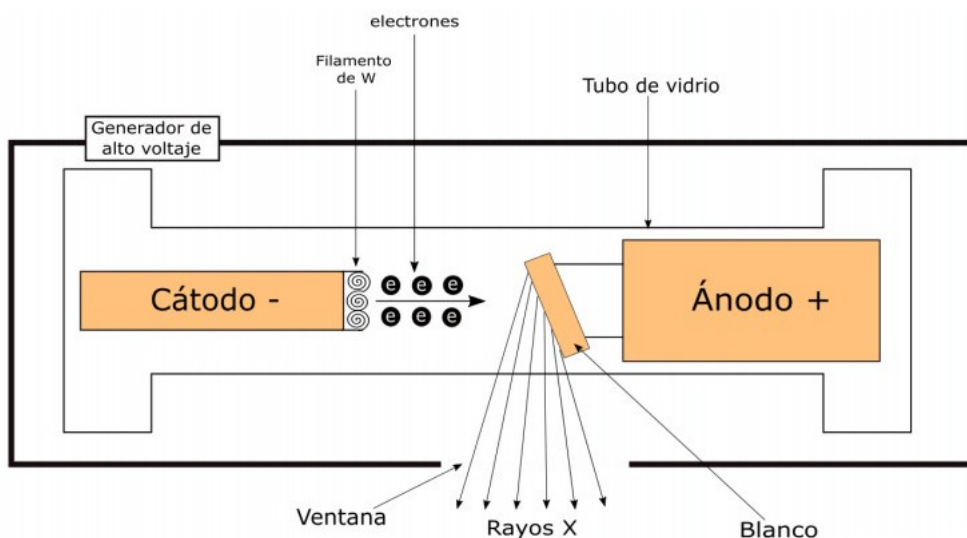


Figura 3: Esquema convencional de un tubo de rayos-X

Al aplicarse el voltaje de aceleración entre el cátodo y el ánodo, usualmente denominado



kilovoltaje (kV), los electrones que han sido desprendidos del cátodo se desplazarán al ánodo acelerándose por el campo eléctrico y alcanzan su velocidad máxima al llegar al ánodo. Estos electrones del proyectil interactúan con los electrones o los núcleos de los átomos del blanco por diversos mecanismos (que los describiremos más adelante) liberando energía en forma de radiación de rayos-X. O en otras palabras, la energía cinética de los electrones provenientes del cátodo es convertida en rayos-X (menos del 1 %) y el resto en calor (mayor al 99 %), por lo tanto la eficiencia de un tubo de rayos-X es muy baja. La energía (poder de penetración) de los rayos-X se expresa en electronvoltios (eV) y es dependiente de la corriente aplicada al filamento del cátodo y de la diferencia de potencial aplicada entre el filamento y el blanco.

2.4. Interacción de rayos-X con la materia

Debido al comportamiento dual de la luz (onda-partícula) la radiación al interactuar con la materia puede exhibir comportamiento ondulatorio y corpuscular. La interacción de las partículas cargadas con la materia y de los efectos que se producen constituyen un extenso campo de estudio como la detección de la radiación, fenómenos físico-químicos que se producen cuando la radiación interactúa con la materia viva, blindajes adecuados contra la radiación o seguridad radiológica, obtención de imágenes diagnósticas para su posterior análisis clínico, entre otras,

Cuando un haz de rayos-X interactúa con la materia es posible que se produzcan los siguientes tres fenómenos:

- *Transmisión:*
- *Absorción:*
- *Dispersión (Scattering):*

A continuación detallamos brevemente en la Figura 4 como interactúan los rayos-X para producir las imágenes médicas y los fenómenos que se producen. Como vimos anteriormente los rayos-X son producidos en un tubo al vacío (1), el haz de rayos-X sale a través de una ventana que contiene un colimador para controlar la anchura del haz con la finalidad de destinar la mayor cantidad de radiación ortogonal al área que se desea radiografiar, este haz se propaga en línea recta en todas las direcciones con igual intensidad, una parte se dispersa en el entorno y otra parte (radiación directa) traspasa el cuerpo del paciente (2). De la radiación que llega al

paciente una parte es absorbida por los tejidos del cuerpo, otra parte es dispersada fuera del cuerpo por el **Efecto Compton**(3). Para disminuir el efecto de radiación dispersa se utiliza un dispositivo denominado rejilla antidifusora que se ubica entre el paciente y la placa radiográfica para absorber la radiación dispersa y mejorar la calidad de la imagen. Por último, otra parte de la energía o radiación electromagnética atraviesa o transmite por el cuerpo con la debida atenuación y logra alcanzar la placa radiográfica formando un patrón que representa las diferentes estructuras del cuerpo (4). La imagen de rayos-X se forma por los fotones transmitidos. Finalmente, revelamos los patrones formados en la placa para obtener la imagen definitiva en escala de grises (5).

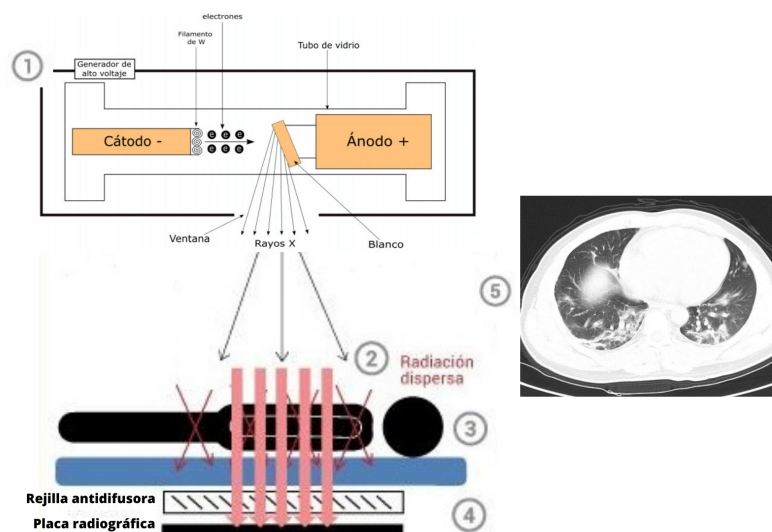


Figura 4: Interacción de rayos-X con un paciente

En radiografía digital, las imágenes se pueden obtener mediante dos métodos: El primero, como lo hemos descrito anteriormente, digitalizando placas o películas (*films*) de rayos-X (radiografía convencional) y la segunda, que los rayos-X que atraviesan al paciente sean capturados por dispositivos electrónicos que convierten a los rayos-X en luz. Y a su vez, esta señal luminosa se transforma en una señal eléctrica, como lo hace la **tomografía axial computarizada** (TAC) que lo describiremos más adelante.

2.5. Píxel y Vóxel

Una imagen digital esta compuesta de un número finito de elementos o celdas, donde cada uno de ellos tiene un valor y una posición particular. Cada celda o “cuadrícula” que forma parte



de una imagen es un **píxel** (acrónimo del inglés *picture element*) que se define como la unidad homogénea en color más pequeña posible. Cuanto mayor sea el número de celdas la imagen se dividirá en mayor número de cuadrículas. Al conjunto de celdas se lo denomina **matriz**. De esta manera si una imagen de TC es de 300 x 200 píxeles sabremos que la imagen estará dividida en 60000 celdas. Vale aclarar que el píxel es una unidad de división sin un tamaño específico y como veremos más adelante solo cuando le asignemos un valor de resolución podemos hablar de un tamaño definido.

Un **vóxel** representa cada uno de los bloques de la imagen que constituyen la matriz a través de la cual es registrada y mostrada una imagen de TC [7], es decir, a cada píxel se lo denota con una profundidad determinada por el espesor con la que se realiza la TC de cada corte. Según el grado de atenuación, a cada píxel se le asocia un color, sea negro, blanco o alguno de la escala de gris.

$$\begin{pmatrix} x_{11} & y_{12} & z_{13} \\ x_{21} & y_{22} & z_{23} \\ x_{31} & y_{32} & z_{33} \\ \vdots & \vdots & \vdots \\ x_{m1} & y_{m2} & z_{m3} \end{pmatrix}$$

$\underbrace{\hspace{1.5cm}}_{\text{eje } x}$
 $\underbrace{\hspace{1.5cm}}_{\text{eje } y}$
 $\underbrace{\hspace{1.5cm}}_I$

donde I es la columna de intensidad.

2.6. Coeficiente de Atenuación lineal (μ)

La atenuación se refiere al hecho que hay menos fotones en el haz emergente que en el haz que entra en el material. Este coeficiente es una relación que compara la pérdida en la intensidad de un haz de energía que pasa a través de un objeto con la distancia que el haz de energía pasa a través del tejido.

Por consiguiente, el haz emitido por el tubo de rayos-X tiene una distribución de intensidad inicial I_0 antes de incidir sobre el paciente, y debido a la absorción y dispersión de fotones que se llevan a cabo al pasar a través de un medio de espesor x , sufriendo una alteración de dicho

haz ya que el medio actúa como filtro del haz puesto que reducirá el valor de la intensidad inicial. Si el haz es monoenergético (o casi), la transmisión de rayos-X a través del paciente es

$$I = I_0 e^{-\mu x} \quad (4)$$

En esta última ecuación se supone que el paciente es un medio homogéneo. Si el haz de rayos-X es interceptado por dos regiones con coeficientes de atenuación μ_1 y μ_2 y de espesores x_1 y x_2 , la transmisión de rayos-X será

$$I = I_0 e^{-(\mu_1 x_1 + \mu_2 x_2)} \quad (5)$$

Y si hay n regiones con diferentes coeficientes de atenuación lineal a lo largo del camino de rayos-x, la transmisión es

$$I = I_0 e^{-(\sum_{i=1}^n \mu_i x_i)} \quad (6)$$

2.7. Unidades Hounsfield (UH)

Son unidades adimensionales utilizadas en tomografía computarizada (TC) que expresan números de TC en forma estándar. Es un número que señala a cada píxel en la TC, por eso a estas unidades también se las conoce como **número de TC** (TCn). Las UH nos brinda la posibilidad de medir el **coeficiente de atenuación** de diferentes tejidos examinados mediante una escala de grises. Su inventor Godfrey Hounsfield lo designó con un rango entre $[-1000, 1000]$ en donde



Figura 5: Unidades Hounsfield (UH), cuyo rango va desde -1000 a 1000

cada uno constituye un nivel diferente de **densidad** (Figura 5). La densidad es medida en la escala de UH, definido aproximadamente de -1000 para aire, 0 UH para agua y de 1000 UH para hueso. Para el caso que nos interesa de los pulmones, consideramos que estos órganos están compuestos fundamentalmente de aire y de “tejido” (incluye sangre, células, agua, etc.), entonces podemos inferir que la densidad en Unidades Hounsfield de los pulmones tendrá un valor negativo que viene de -1000 (UH de aire) y del valor del tejido (0 UH si consideramos el “tejido” como agua) de la ROI analizada.



Los números de TC de los tejidos son variables debido a la heterogeneidad de esos mismos tejidos y la variación en el coeficiente de atenuación de cada tejido en relación al agua, que depende del kilovoltaje (kV) y de la filtración del haz de rayos-X [5]. Algunos valores representativos de tejidos los podemos ver en la Tabla 2 y en la Figura 6. Los coeficientes de la figura fueron analizados utilizando el software *RadiAnt DICOM*.

Tejido	UH (CT number)
Aire	-1000
Pulmón	-500
Grasa	-50 a -100
Agua	0
Músculo	10 a 40
Riñon	30
Sangre	40
Hueso	1000

Tabla 2: Rango aproximado de Unidades Hounsfield para diferentes tejidos

2.7.1. Fórmula

A estas unidades se las representa mediante

$$UH = 1000 \frac{\mu_{tejido} - \mu_{agua}}{\mu_{agua} - \mu_{aire}} \quad (7)$$

donde μ_{agua} es el coeficiente de atenuación del agua destilada, μ_{tejido} es el coeficiente de atenuación del tejido de nuestro interés y μ_{aire} es el coeficiente de atenuación del aire que tiene un valor de cero ($\mu_{aire} = 0$), por lo que la expresión anterior quedaría de la siguiente forma:

$$UH = 1000 \frac{\mu_{tejido} - \mu_{agua}}{\mu_{agua}} \quad (8)$$

2.8. Contraste de una imagen

El contraste es una de las mayores particularidades de una imagen que describe que tan bien la imagen distingue características sutiles en el objeto (paciente). En diagnóstico por imagen, el contraste de una imagen es el resultado de interacciones complejas entre atributos anatómicos y fisiológicos de la región de interés [8].

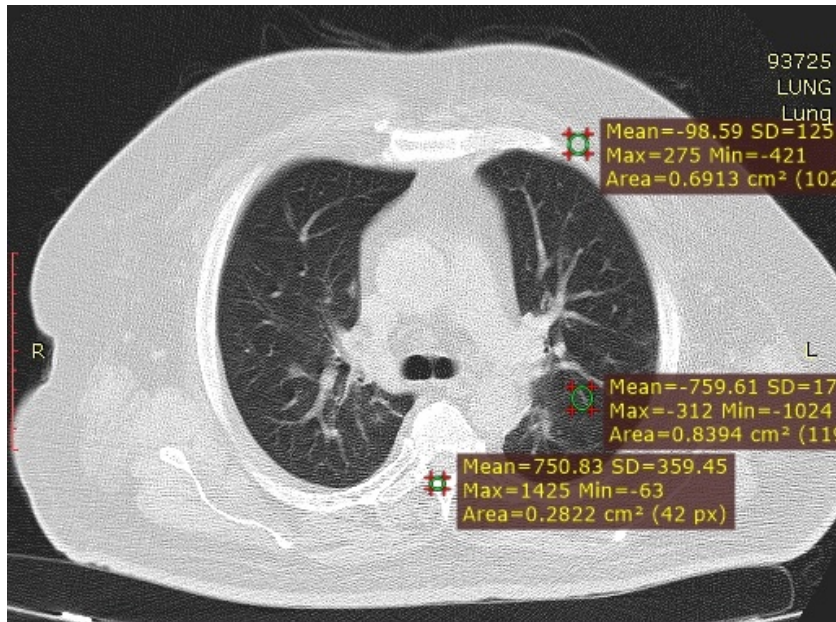


Figura 6: Valores de UH del pulmón, grasa (tejido adiposo) y hueso de una TC de tórax de paciente con COVID-19

2.9. Tomografía Lineal

En términos generales tomografía hace referencia a la “vista de corte” de un paciente, aunque también es común relacionarla como “imagen seccional”. La tomografía es una técnica de obtención de imágenes por rayos-X (RX) que genera vistas seccionales de los pacientes en un plano paralelo a la superficie de la mesa de exploración sobre la cual el paciente está acostado [5].

2.10. Tomografía Computarizada (TC)

La TC es una herramienta de diagnóstico que produce imágenes en secciones trans-axiales, es decir, perpendiculares al eje de rotación de la fuente de rayos-X sobre el cuerpo [5]. Tanto la fuente de rayos-X que se proyecta a un paciente desde diferentes ángulos y el detector de rayos-X rotan alrededor del cuerpo, generando así señales que son procesadas por una computadora para formar las imágenes transversales que representan un “corte” del paciente. El conjunto de tales cortes constituyen una representación tridimensional (3D) de alguna parte del interior del cuerpo.

La diferencia principal de una TC y una radiografía convencional es que una TC utiliza una fuente de escáner motorizada (en movimiento) que rota alrededor de una estructura en forma de anillo o dona llamada **gantry**, mientras que en la radiografía convencional su funcionamiento se basa en un tubo fijo de rayos-X. Además, las imágenes o cortes obtenidos en TC proveen una información mas detallada que las radiologías tradicionales [9]. Así, el proceso de TC consiste en la emisión de rayos X que son atenuados por los tejidos y estructuras que están en el paso de los rayos X a través del paciente y son recolectados en el extremo opuesto por los correspondientes detectores en el gantry. Este procedimiento se va repitiendo mientras la fuente gira.

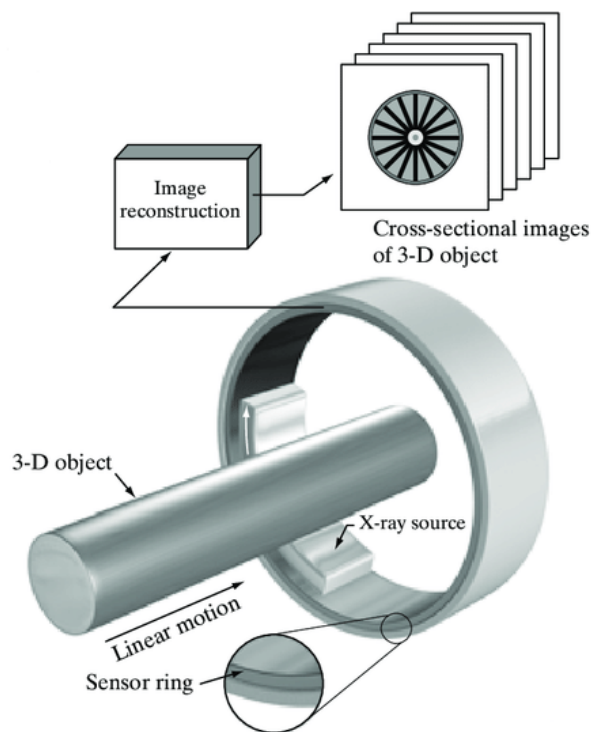


Figura 7: Adquisición de imágenes usando una fuente giratoria de rayos-X

La imagen capturada por un tomógrafo no es más que una matriz (compuesta por filas y columnas) donde cada elemento de dicha matriz es un píxel, por ejemplo una matriz de 256 x 256 píxeles, o también es común que la imagen de TC se calcule en matrices de 512 x 512 , 1024 x 1024, etc.



2.11. Imágenes Digitales

Los procesos de obtención, recepción, almacenamiento, visualización y reporte de resultados de las imágenes médicas se lo debemos al sistema de archivos y comunicación de imágenes (PACs, por las siglas anglosajonas de *Picture Archiving and Communications Systems*) y la especialidad que se encarga de todo este estudio es la **Radiología**. Esta rama de la ciencia brinda información relevante de las estructuras y funciones de órganos o tejidos analizados. Las imágenes pueden ser visualizadas desde cualquier pantalla de un ordenador y también en formato impreso, por ejemplo en una lámina de poliéster.

Consideramos una imagen como una función bidimensional $f(x, y)$, donde x y y son coordenadas espaciales y la amplitud de f en cualquier par de coordenadas (x, y) es conocido como **intensidad** o **nivel de gris** de la imagen en ese punto. Cuando los valores de x, y y la intensidad f son cantidades finitas y discretas la imagen es denominada **imagen digital** [10]. La función $f(x, y)$ viene determinada por dos componentes: **iluminación** y **reflectancia** y son denotadas por $i(x, y)$ y $r(x, y)$, respectivamente. En el caso de imágenes producidas mediante transmisión de ondas electromagnéticas a través de un medio (como una tomografía computarizada de tórax) al referirnos a *iluminación* corresponderá a la radiación electromagnética (rayos-X) y en lugar de *reflectancia* lo más conveniente sería hablar de **transmisividad**. Entonces, $f(x, y)$ viene dado por

$$f(x, y) = i(x, y)r(x, y) \quad (9)$$

donde i es la cantidad de iluminación (energía) de la fuente que incide en el objeto y r es la cantidad de iluminación transmitida por el objetos. La naturaleza de i está determinada por la fuente de iluminación, donde

$$0 \leq i(x, y) < \infty \quad (10)$$

y la transmisión está limitada por 0 (transmisión total) y 1 (absorción total)¹

$$0 \leq r(x, y) \leq 1 \quad (11)$$

por consiguiente, el valor de las coordenadas de $f(x, y)$ es una cantidad positiva y finita [10],

¹Cuando se trata de imágenes que no son formadas por transmisión de luz se la puede caracterizar por **reflectancia** que está limitada por 0 (absorción total) y 1 (reflectancia total).



esto es,

$$0 \leq f(x, y) < \infty \quad (12)$$

2.11.1. Intensidad

Los huesos, algunos tumores y otra materia densa aparecen en color blanco o claro en virtud de que absorben la radiación. Por este motivo, el análisis de la distribución de intensidades presentes en una TC nos proporciona una información muy valiosa, ya que cada imagen está compuesta por un conjunto de píxeles en escala de grises que representan las diferentes densidades de los tejidos en la TC.

Las ondas electromagnéticas viajan en líneas rectas llamados rayos (representan la dirección en la que la radiación se propaga) que irradian en todas las direcciones desde una fuente puntual. Al conjunto colimado de rayos se le conoce como **haz** [5]. La intensidad de los rayos-X se modifica por absorción a medida que atraviesan el paciente [10]. La intensidad (nivel de grises) de una imagen monocromática se denota por

$$\mathcal{I} = f(x, y) \quad (13)$$

y en teoría el rango de \mathcal{I} es igual a la expresión (12), pero en la práctica comprende el siguiente rango

$$L_{min} \leq \mathcal{I} < L_{max} \quad (14)$$

donde $L_{min} = i_{min}r_{min}$ y $L_{max} = i_{max}r_{max}$. Y el rango $[L_{min}, L_{max}]$ se le denomina **escala de intensidad** (o escala de grises)[10]. Es muy frecuente cambiar este intervalo numéricamente al intervalo $[0, 1]$, donde $\mathcal{I} = 0$ se considera negro y $\mathcal{I} = 1$ representa blanco en la escala. Y todos los valores intermedios son tonos de gris que varían del negro al blanco. En la presente investigación usaremos esta última escala.

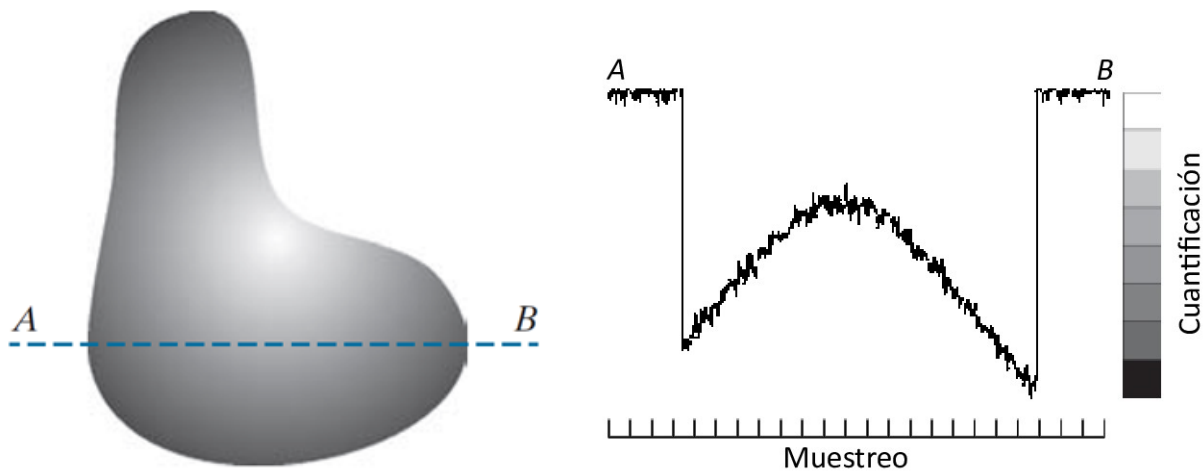
2.11.2. Muestreo y Cuantificación de Imágenes

Para crear una imagen digital se necesita convertir los datos obtenidos continuamente a un formato digital. Esto se consigue mediante dos procesos: **muestreo** y **cuantificación**. En la figura 8a se muestra una imagen continua que se desea convertir en formato digital. Una imagen puede ser continua con respecto a las coordenadas x y y y también en amplitud. Es decir, para



digitalizarla se tiene que muestrear la función tanto en coordenadas como amplitud. La digitalización de los valores de las coordenadas se lo conoce como muestreo (en inglés *sampling*) y la digitalización de los valores de amplitud (nivel de intensidad) se lo llama cuantificación (en inglés *quantization*).

La función de la Figura 8b es una gráfica de los valores de amplitud o niveles de intensidad de la imagen a lo largo de la línea AB en la Figura 8a. Se aprecia notables variaciones en los extremos debido al ruido de la imagen. La barra vertical representada en la Figura 8b representa la escala de intensidad (para este ejemplo esta dividido en 8 intervalos discretos) que van del negro al blanco.



(a) Imagen continua con una línea AB que muestra las variaciones de intensidad

(b) Muestreo y cuantificación

Figura 8: Muestreo y cuantificación de una imagen

En la práctica, el método de muestreo está determinado por la disposición de sensores utilizados para generar una imagen [10]. Por ejemplo, para una imagen generada por un único sensor en movimiento la salida del sensor se cuantifica de la manera que se lo describió en la figura 8. Cuando se utiliza una tira o banda de sensores la cantidad de sensores en la tira define las muestras en la imagen resultante en una dirección y el movimiento mecánico establece el número de muestras en otra dirección [10]. El proceso de generación de una imagen digital se completa con la cuantificación de las salidas del sensor.

Cuando se utiliza un conjunto (matriz) de sensores (en el caso de un tomógrafo tiene una tira

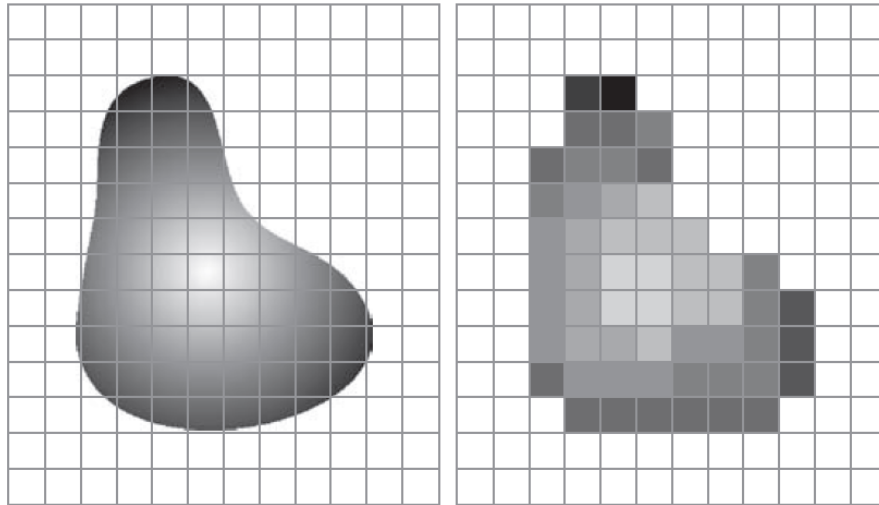


Figura 9: A la izquierda, imagen continua proyectada sobre una matriz de sensores. A la derecha, la imagen luego del muestreo y cuantificación.

circular de sensores fijos) no es necesario el movimiento mecánico. En la figura 9 a la izquierda se muestra una imagen proyectada sobre un arreglo de sensores y a su lado se muestra la imagen después del muestreo y la cuantificación. La calidad de una imagen digital es determinado en gran medida por el número de muestras y los niveles de intensidad discretos usados en el muestreo y la cuantificación.

2.11.3. Representación digital de imágenes

La presentación de una imagen tradicional se realiza en un mapa bidimensional o retícula de cuatro lados, iguales de dos a dos. Las celdas individuales con la que está dividida la retícula (matriz) es lo que conocemos como píxeles. Para conocer la dimensión de una imagen se nombra siempre primero el número de píxeles contenidos en el ancho (que en este estudio lo simbolizaremos con la letra M) y luego el número de píxeles que hay en el alto (N) de dicho mapa, es decir, el tamaño de una imagen será $M \times N$ (# píxeles ancho x # píxeles alto).

Supongamos que deseemos muestrear una imagen continua en una imagen digital, $f(x, y)$ contiene M filas y N columnas, donde (x, y) son coordenadas discretas. Por conveniencia denotamos: $x = 0, 1, 2, \dots, M - 1$ y $y = 0, 1, 2, \dots, N - 1$. Por ejemplo, el valor de una imagen digital en el origen es $f(0, 0)$ y su valor en la siguiente coordenada a lo largo de la primera



fila es $f(0, 1)$. En general, los valores de una imagen digital en cualquier coordenada (x, y) es denotado por $f(x, y)$, donde x y y son enteros. Es decir, cuando necesitemos referirnos a una específica coordenada (i, j) usamos la notación $f(i, j)$, donde los argumentos son enteros.

$f(x, y)$ se puede representar en una matriz de números reales. Esta es la representación usada para el procesamiento por computadora. En ecuación, la representación de una matriz $M \times N$ es

$$f(x, y) = \begin{bmatrix} f(0, 0) & f(0, 1) & \cdots & f(0, N - 1) \\ f(1, 0) & f(1, 1) & \cdots & f(1, N - 1) \\ \vdots & \vdots & \ddots & \vdots \\ f(M - 1, 0) & f(M - 1, 1) & \cdots & f(M - 1, N - 1) \end{bmatrix} \quad (15)$$

Cada elemento de la matriz es llamado *picture element* o *píxel*. El centro de una imagen digital $M \times N$ con el origen en $(0, 0)$ y rango $(M - 1, N - 1)$ (como se muestra en la Figura 10) se obtiene dividiendo M y N por 2 y redondeando hacia abajo al entero más cercano. Muchas veces esta operación es denotada usando el operador *floor*, entonces $(x_c, y_c) = (\text{floor}(M/2), \text{floor}(N/2))$. Por ejemplo, el centro de una imagen de 1023×1024 es $(511, 512)$.

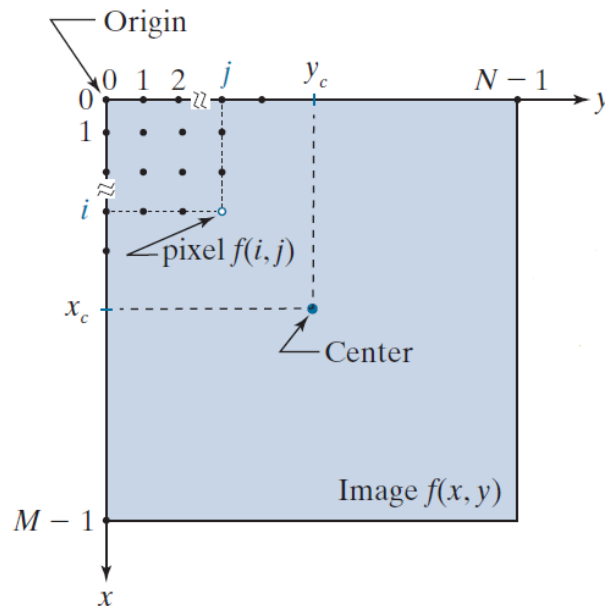
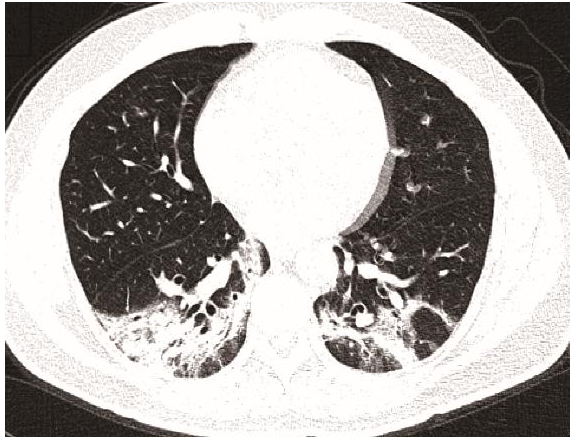


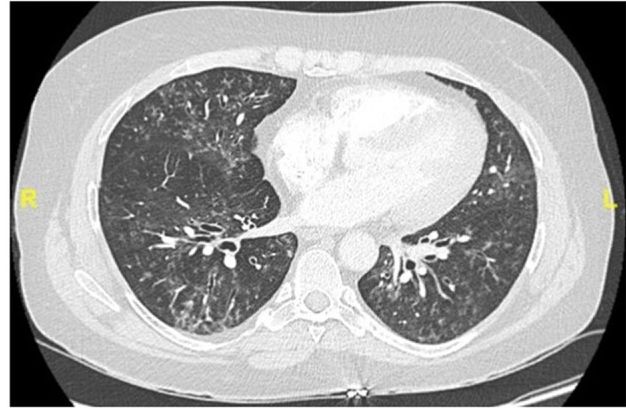
Figura 10: Convención de coordenadas usadas para representar una imagen digital.

Cada píxel tiene un valor asignado que está relacionado con la intensidad de la señal. El tamaño de la matriz (es decir el número de píxeles), el tamaño de los píxeles y el campo de visión

(el área que queremos estudiar) están relacionados. Por ejemplo, la mayoría de las imágenes usadas en TC se muestran en una matriz de 512 x 512 para un total de 262144 píxeles y cubren un campo de visión que puede ser de 350 mm x 350 mm. Entonces, para este caso el tamaño de cada píxel es de 350 mm/512 que es aproximadamente 0,7 mm.



(a) Imagen original de TC paciente positivo para COVID-19



(b) Imagen de TC para paciente negativo para COVID-19

Figura 11: Imágenes de TC de tórax adquiridas durante la pandemia por COVID-19

Como vimos anteriormente, a una imagen digital se le puede asociar como una matriz de números, de esta forma se pueden realizar ciertas operaciones que se calculan con las matrices numéricas (suma, resta, multiplicación, división, etc). Igualmente, es posible realizar operaciones relacionadas así como: $<$, $>$, \leq , \geq . Además, dentro de los lenguajes de programación se pueden manipular estos valores usando diferentes ciclos como: *if*, *while*, *for*, *do*, etc.

2.11.4. Resolución

Al hablar de resolución nos referimos al número de píxeles o puntos que componen una imagen por pulgada. Así pues, se puede medir tanto en **píxeles por pulgada** (ppp), (o en inglés ppi, siglas de *pixels per inch*) y también en *puntos por pulgada* (dpi, siglas de *dots per inch*). Aunque lo habitual es referirse como número de *píxeles por pulgada* para una imagen cuando es reproducida en una pantalla de una computadora y como número de *puntos por pulgada* para una imagen cuando es impresa. Por ejemplo, en una imagen con una resolución de 400 ppp, cada pulgada de la imagen posee 400 píxeles. O en otras palabras, cada píxel de la imagen



ocupa 1/400 de pulgada. Si se divide el número de píxeles del ancho y del alto de la imagen por la resolución se obtiene el tamaño de la imagen ya sea en una pantalla o impresa.

Evidentemente, el tamaño de la imagen está relacionado con la resolución, lo que condiciona la calidad de la imagen o la capacidad de distinguir detalles finos. Mientras más píxeles tenga una imagen o mayor sea su resolución, menor será el tamaño del píxel y se apreciará mejor la cantidad de detalles. La *resolución espacial* es la dimensión del tamaño del píxel que representa el área del píxel dentro de la imagen. Por ejemplo, la imagen de TC de la Figura 11b tiene un tamaño de 630 x 416 píxeles y una resolución de 96 ppp. Entonces, cada píxel que compone la imagen ocupa 1/96 de pulgada, y el tamaño físico de la imagen será $630/96 \approx 4.33$ pulgada y $416/96 \approx 6.56$ pulgada para el ancho y alto, respectivamente. Hay que considerar que cada imagen de TC pueden tener un tamaño y resolución característico.

2.11.5. Profundidad de color

La profundidad de color hace referencia al número de **bits** necesarios para representar o codificar la información del color de cada píxel en una imagen. Un bit, del acrónimo *binary digit*, que simboliza la unidad de información más pequeña posible, es un dígito del sistema de numeración binario que se representa con valores de 0 y 1. Por lo tanto, en una imagen de **1 bit** cada píxel solo puede tener uno de dos valores ya sea el cero o el uno; **2 bits** abarcará 4 valores que serán 00, 01, 10, 11; para **3 bits** cada píxel puede tener 8 valores diferentes (000, 001, 010, 011, 100, 101, 110, 111); y así sucesivamente.

El número de niveles de intensidad L es una potencia entera de 2, esto es

$$L = 2^n \quad (16)$$

donde n es un entero. Entonces, para apreciar la cantidad de colores posibles que contendrá una imagen se debe elevar el 2 a una potencia de n bits, es decir, 2^n . Una imagen de escala de grises es de 8 bits, esto es,

$$L = 2^n = 2^8 = 256,$$

lo que significa que cada píxel puede tener un valor entre 256. Por consiguiente, en la escala de grises un píxel puede ser negro, blanco o tomar un tono gris entre las 254 posibilidades



restantes. Por conveniencia el rango de valores se escala de la siguiente manera

$$[0, L - 1], \quad (17)$$

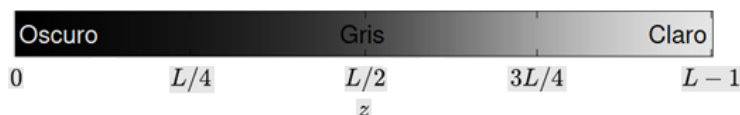


Figura 12: Rango de valores en escala de grises

que para el caso de 8 bits es igual a $[0, 255]$. Al grupo de 8 bits comúnmente se lo conoce como **byte**, por lo tanto una imagen de 256 colores es una imagen de 1 byte.

2.11.6. Píxeles vecinos

Un píxel p de coordenadas (x, y) tiene dos vecinos horizontales y dos vecinos verticales con coordenadas

$$(x + 1, y), (x - 1, y), (x, y + 1), (x, y - 1)$$

Este conjunto de píxeles, llamado los *4 vecinos de p* (ver figura 13a), se denota por $N_4(p)$.

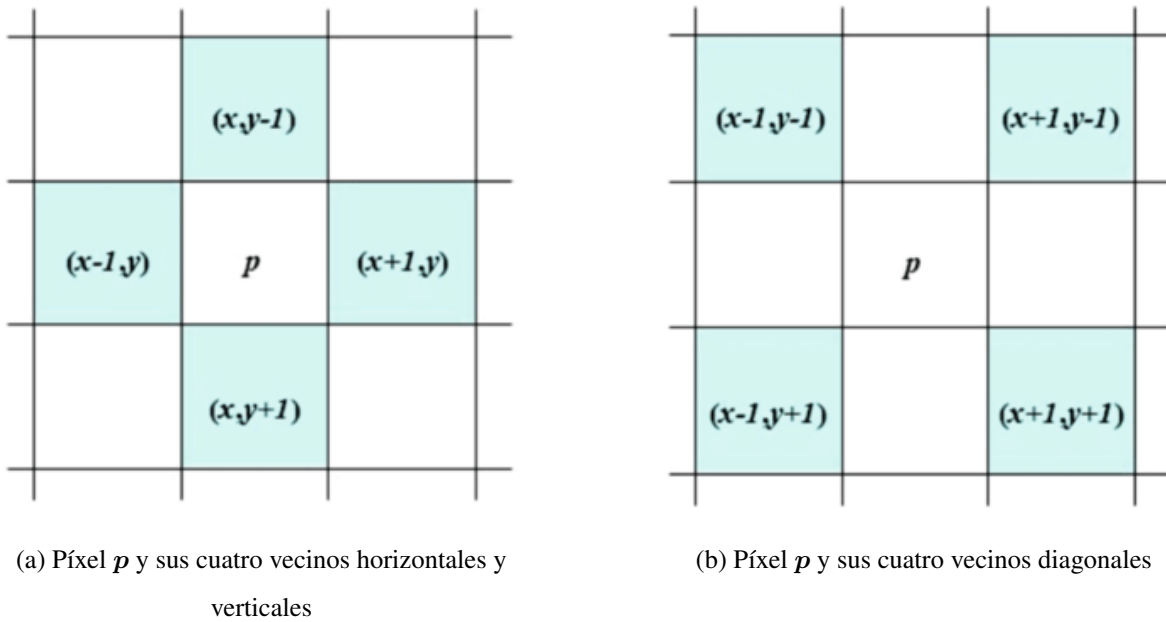
Asimismo, los 4 vecinos diagonales de p tienen coordenadas (ver figura 13b)

$$(x + 1, y + 1), (x - 1, y - 1), (x - 1, y + 1), (x + 1, y - 1)$$

y son denotados por $N_D(p)$. Este conjunto de ubicaciones de los vecinos de un punto p se denomina *vecindad* de p .

Profundidad de color	Colores posibles	Característica
1 bit por píxel	2	Modo monocromático (Blanco y negro)
2 bit por píxel	4	Blanco y negro
4 bit por píxel	16	Modo escala de grises
8 bit por píxel	256	Modo escala de grises, admite formatos GIF y PNG
16 bit por píxel	65536	High color

Tabla 3: Distintos valores de bits para representar el color píxel en una imagen.

Figura 13: Vecinos de un píxel p

2.12. Digitalización y Análisis de Imágenes

Uno de los principales desafíos que tiene un especialista en imágenes o un radiólogo es el análisis de imágenes obtenidas por las diferentes técnicas conocidas tales como rayos X, tomografía computarizada (TC), resonancia magnética (RM), tomografía por emisión de positrones (PET), entre otras. Actualmente, estas imágenes pueden ser adquiridas por medio de detectores que convierten, por ejemplo, los rayos X en imágenes digitales.

Para analizar imágenes médicas como radiografías o TC es conveniente hablar de lo que se conoce como *radiografía modo máscara* un recurso muy beneficioso en la sustracción de imágenes. Se considera una diferencia o sustracción de imágenes, así

$$g(x, y) = f(x, y) - h(x, y) \quad (18)$$

En este caso $h(x, y)$ representa la **máscara**. El efecto de restar la máscara de cada imagen $f(x, y)$ es que las áreas que son diferentes entre $f(x, y)$ y $h(x, y)$, aparecen en la imagen de salida, $g(x, y)$ como detalle mejorado [10]. Una importante aplicación de la *multiplicación* de imágenes es en la selección de las regiones de interés (en inglés *masking*). Este proceso consiste en multiplicar una imagen por una imagen de máscara que puede tener, por ejemplo, 1's en la ROI y 0's en otros lugares (ecuación 19). Puede haber más de una ROI en la imagen de una

máscara y la forma de la ROI puede ser arbitraria [10].

$$g(x, y) = f(x, y) h(x, y) \quad (19)$$

Todos los programas que se utilizaron en esta investigación fueron escritos en el lenguaje de programación **C** y compilados en el *framework* de análisis de datos conocido como *ROOT*. Estos algoritmos se utilizan para codificar los niveles de intensidad en colores para facilitar la interpretación humana de las imágenes de TC y en procedimientos para extraer información de una imagen para su correcto procesamiento.



Capítulo 3

3. Base de datos

La mayoría de las imágenes de nuestro repositorio contienen la siguiente particularidad: En lo que se refiere al ruido tenemos dos tonos muy bien definidos, uno muy oscuro o negro y otro muy claro o blanco. Debido a lo anterior hemos separado el ruido en “ruido oscuro” y “ruido claro” para las regiones “negras” y “blancas”, respectivamente. En la tabla 11 (Ver Apéndice) apreciamos el valor del promedio de la media y del promedio de los sigmas para el ruido oscuro de una muestra de 30 pacientes diagnosticados con COVID-19. (El rango de μ es de 0 a 1). En la tabla 12 (Ver Apéndice) apreciamos el valor del promedio de la media y del promedio de los sigmas para el ruido claro de una muestra de 30 pacientes. Se ha organizado en algunos grupos las imágenes para obtener patrones comunes que serán de gran utilidad con el objetivo de la clasificación de las imágenes. Tenemos imágenes con opacidades “circulares”, regiones “faltantes”, regiones “negras” y regiones “opacas”. Mostraremos unos pocos ejemplos de cada uno de estos grupos.

3.0.1. Imágenes con opacidades circulares

Observamos que la distribución del histograma de este paciente con COVID-19 con opacidades circulares está concentrada en los valores **altos** de intensidad esto a su vez se ve reflejado en la imagen correspondiente, la cual es predominantemente **clara**, y con poco contraste.

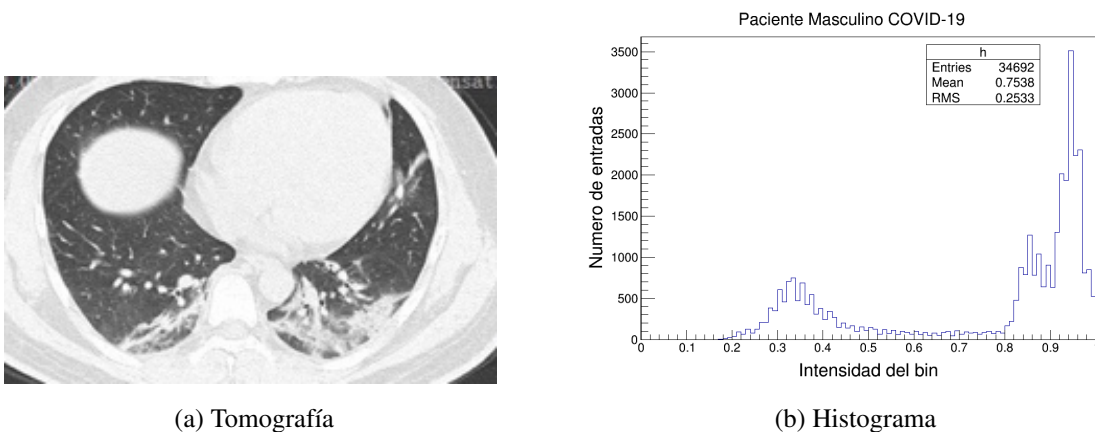
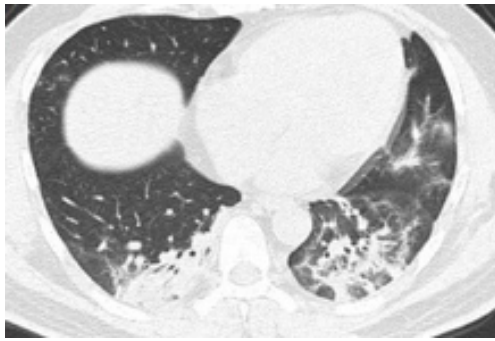
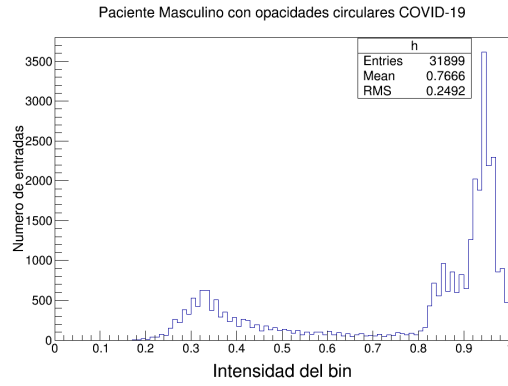


Figura 14: Paciente Masculino con opacidades circulares 1

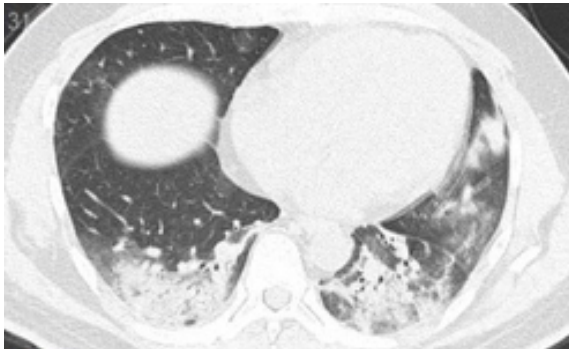


(a) Tomografía

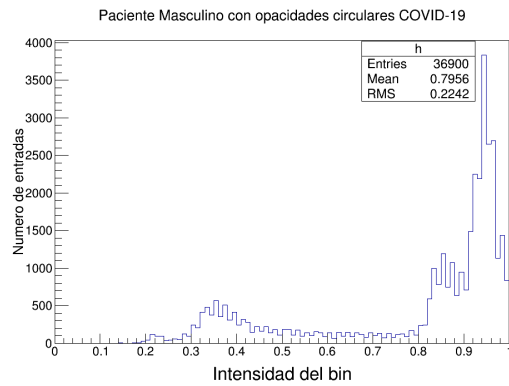


(b) Histograma

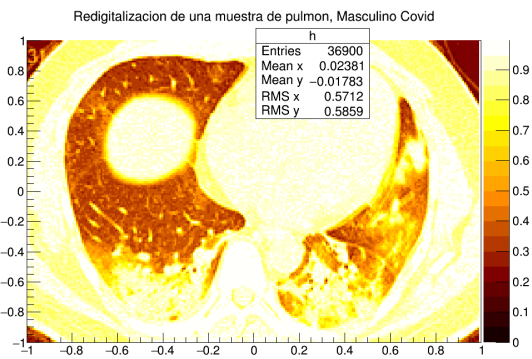
Figura 15: Paciente Masculino con opacidad circular 2



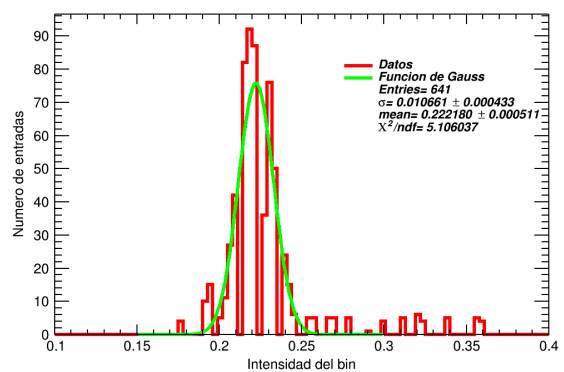
(a) Tomografía



(b) Histograma

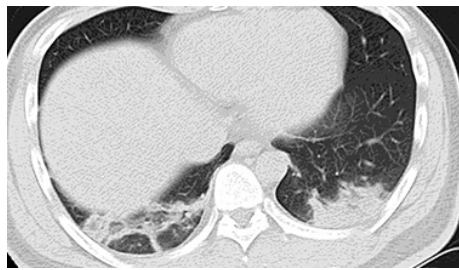


(c) Mapa Calor

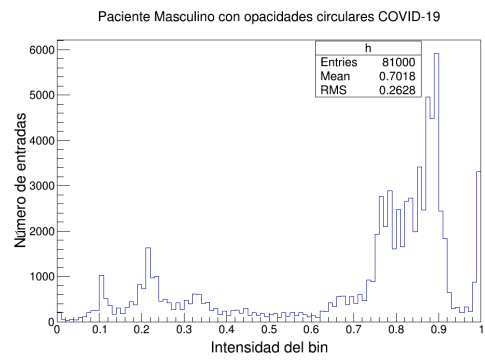


(d) Ruido oscuro

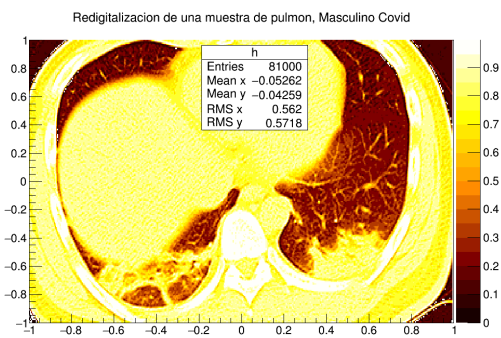
Figura 16: Paciente Masculino con opacidad circular 3



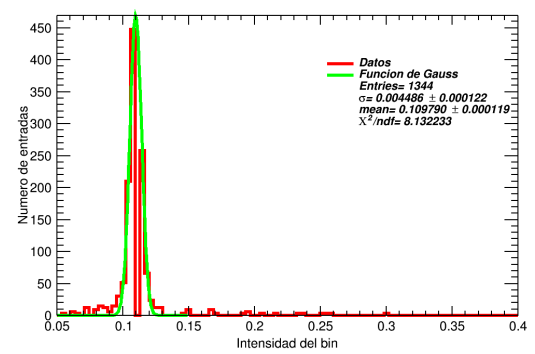
(a) Tomografía



(b) Histograma



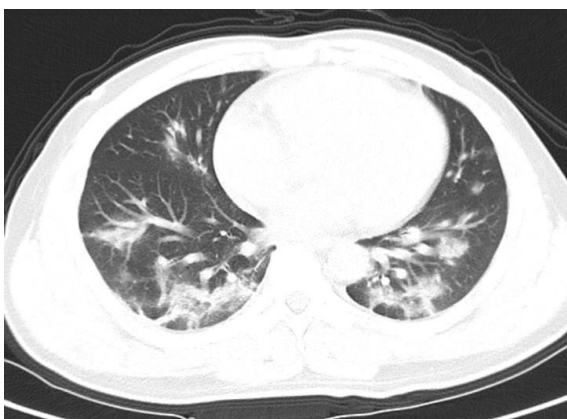
(c) Mapa Calor



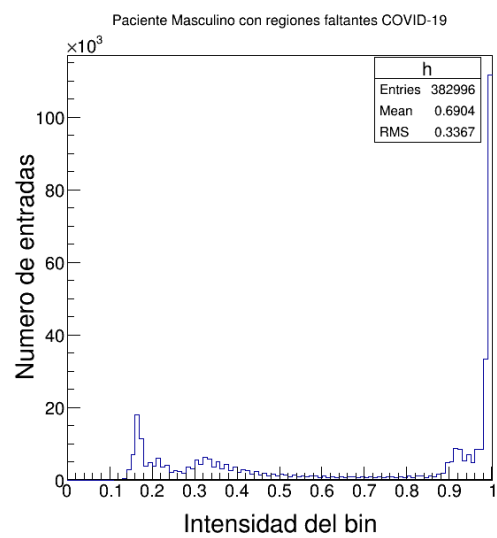
(d) Histograma Ruido

Figura 17: Paciente Masculino con opacidad circular 4

3.0.2. Imágenes con regiones faltantes



(a) Tomografía

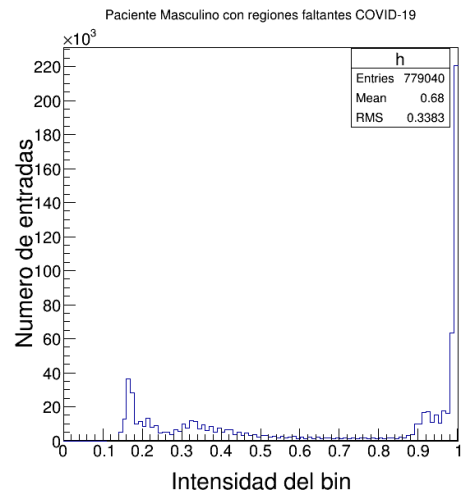


(b) Histograma

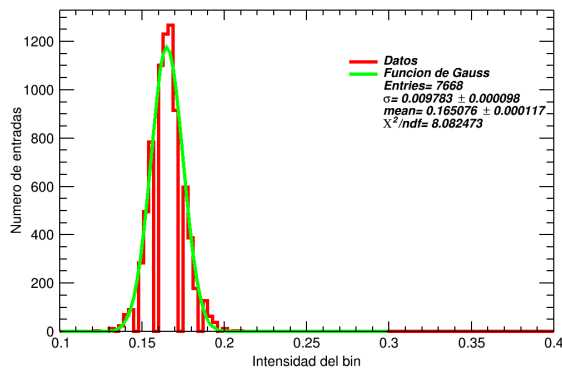
Figura 18: Paciente Masculino con áreas blancas



(a) Tomografía

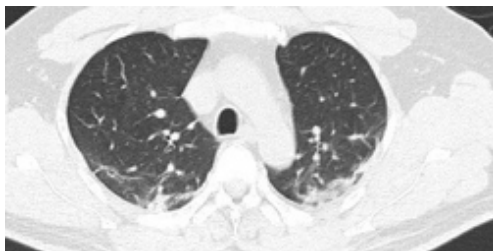


(b) Histograma

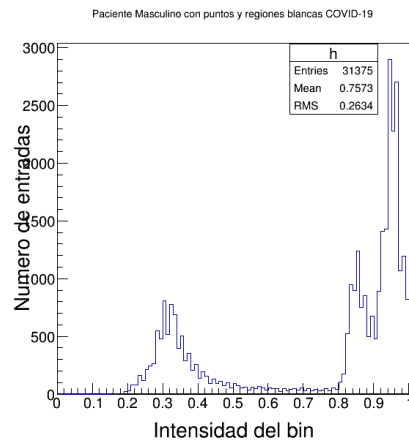


(c) Ruido oscuro

Figura 19: TC de paciente Masculino con regiones faltantes



(a) Tomografía

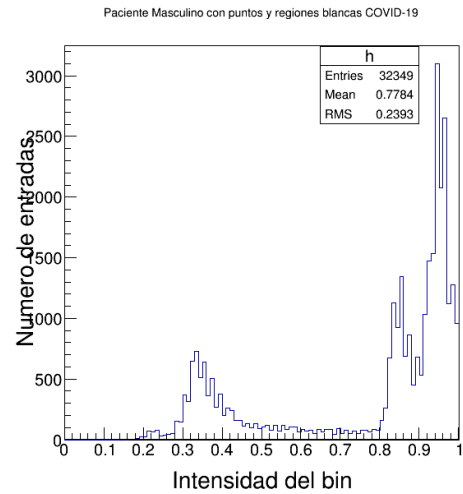


(b) Histograma

Figura 20: Paciente Masculino 20



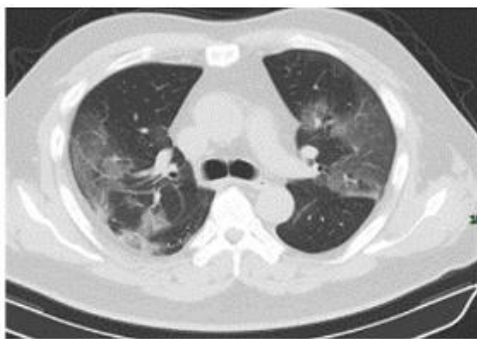
(a) Tomografía



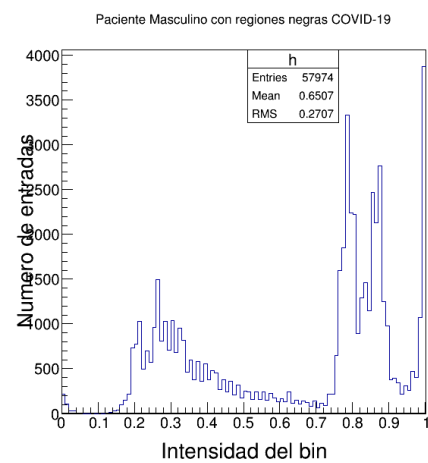
(b) Histograma

Figura 21: Paciente Masculino 21

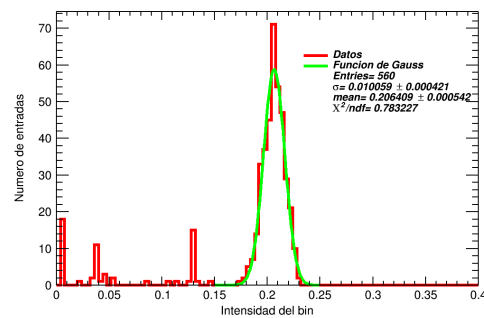
3.0.3. Imágenes con regiones negras



(a) Tomografía



(b) Histograma



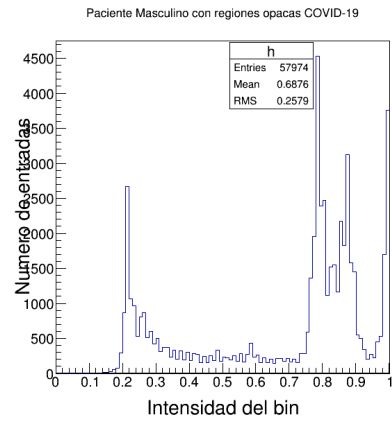
(c) Ruido oscuro

Figura 22: TC paciente masculino con regiones negras

3.0.4. Regiones Opacas



(a) Tomografía

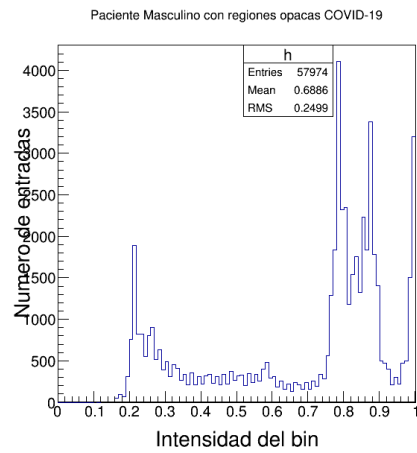


(b) Histograma

Figura 23: TC paciente Masculino 23



(a) Tomografía

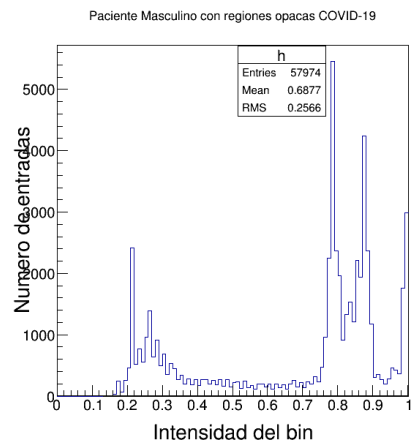


(b) Histograma

Figura 24: TC paciente Masculino 24



(a) Tomografía



(b) Histograma

Figura 25: TC paciente Masculino 25



3.1. Imágenes No COVID-19

A continuación se muestra algunas de las imágenes del repositorio que se usaron para el análisis de datos. Junto a cada imagen está su respectivo histograma de intensidades.

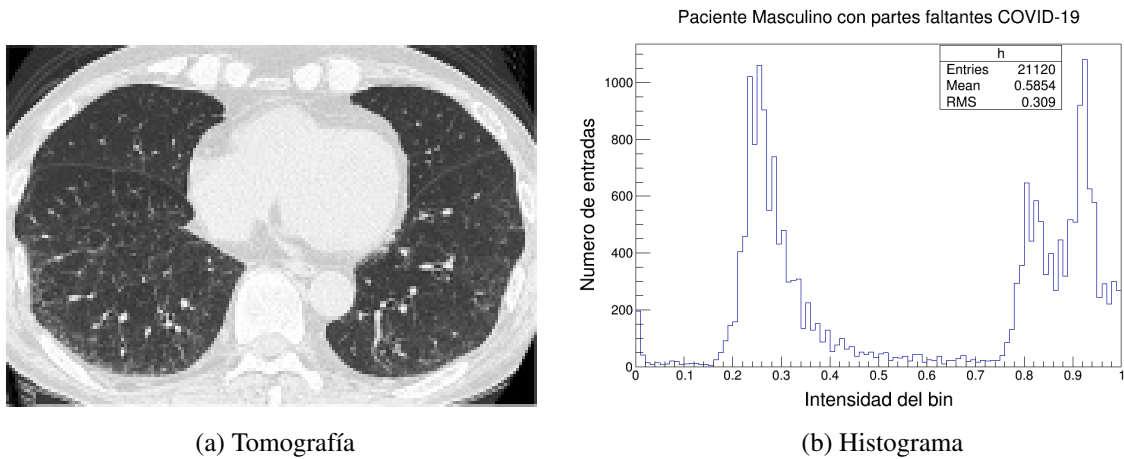


Figura 26: Paciente No-Covid 26

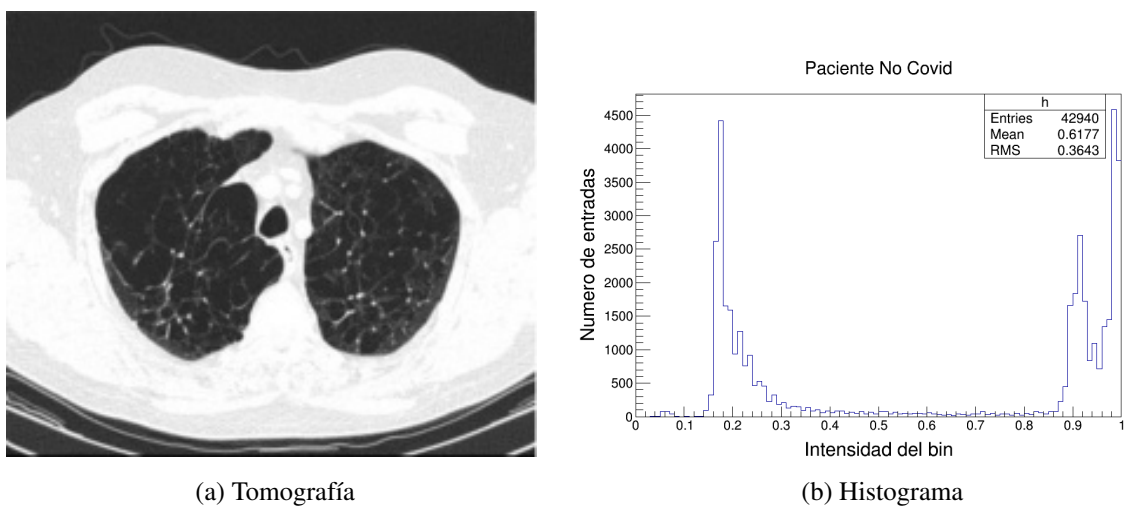
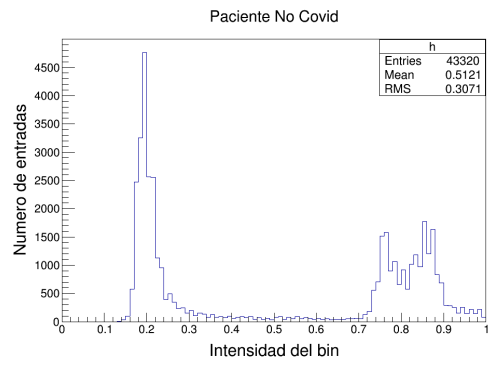


Figura 27: Paciente No-Covid 27



(a) Tomografía

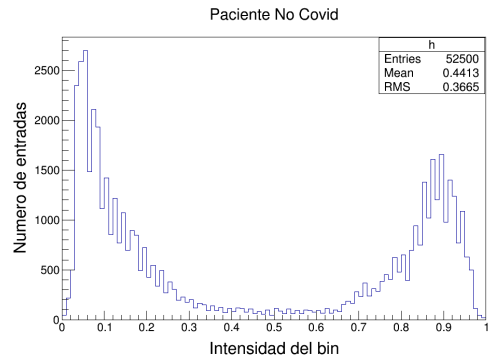


(b) Histograma

Figura 28: Paciente No-Covid 28



(a) Tomografía

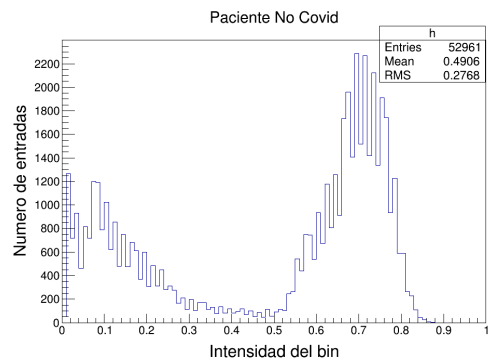


(b) Histograma

Figura 29: Paciente No-Covid 29



(a) Tomografía



(b) Histograma

Figura 30: Paciente No-Covid 30



Capítulo 4

4. METODOLOGÍA

4.1. Procesamiento de Imágenes

El procesamiento de imágenes se utiliza para emitir diagnósticos, estimar los cambios que se vayan produciendo en algún órgano o tejido en análisis, tomar una decisión acerca de un tratamiento, entre otros. Debido al gran avance en las técnicas de adquisición de imágenes y el alto desarrollo de los equipos médicos es idóneo tener presente como hacer un proceso con gran exactitud para el éxito de un diagnóstico clínico.

En la presente investigación el procesamiento de las imágenes de TC empieza mediante unos algoritmos escritos en el lenguaje de programación C que transforman la imagen original en otra para resaltar ciertas *regiones de interés* y también para atenuar o eliminar áreas de la imagen que son irrelevantes (ruido) para nuestro estudio. Como se explicará más adelante en la sección 4, este proceso comprende la transformación de tono de gris a un modelo de color elegido, eliminación del ruido y segmentación de los lóbulos pulmonares.

Además de los algoritmos de C que son ejecutados en el entorno de trabajo ROOT, en este trabajo para el procesamiento digital de las imágenes de TC se utilizó ocasionalmente el programa **ImageJ** que nos permite visualizar, editar, procesar, analizar, transformar imágenes, resaltar detalles y ocultar otros, etc. Puede leer formatos PNG y JPEG en las que están archivadas nuestras TC, por otro lado, esta herramienta es muy útil para operaciones de imágenes “muy grandes” puesto que este programa es muy rápido y puede ejecutar varias imágenes a la vez en paralelo.

4.2. Generación de la matriz

Las imágenes de TC de la base de datos vienen en formato de archivo *png*, esta extensión permite comprimir archivos sin pérdida de información. Como punto de partida para el posterior análisis de datos se *codifican* las imágenes como matrices. Por medio de un programa al que



se le ha nombrado *ImageToHisto.C* (escrito en C) creamos una matriz tridimensional (x, y, z) para cada imagen de TC seleccionada, los dos primeros números representan las coordenadas de la posición y z representa la **intensidad** de un píxel. Cada matriz formada consta de M filas por 3 columnas ($\mathcal{M}_{M \times 3}$), siendo M el número de píxeles con la que esta formada cada una de las imágenes.

La función que nos permite la lectura de las imágenes en el código es *TASImage image* (“*Imagen1.png*”) y el parámetro que despliega la matriz es *std::ofstream outFile*(“*Imagen1.dat*”). Cada una de estas matrices generadas las almacenamos en un archivo de texto diferente, cuya extensión puede ser *.txt* o *.dat* que son compatibles con los programas empleados.

4.3. Escala de grises

Como se puede observar en el espectro EM de la figura 2 hay una banda visible para el ojo humano que abarca longitudes de onda entre aproximadamente $0.43 \mu\text{m}$ (violeta) a alrededor de $0.79 \mu\text{m}$ (rojo) [10]m, donde este espectro de **colores** se divide en 7 regiones: violeta, azul, cian, verde, amarillo, naranja y rojo. Y a la luz que carece de color se le denomina **monocromática** (o acromática) [10] y su principal atributo es su **intensidad**. Y esta intensidad de la luz monocromática varía desde el negro, tonos grises hasta llegar al blanco y se usa el término **nivel de grises** para expresar la intensidad monocromática y el rango de valores desde el negro al blanco se lo denomina **escala de grises** y a las imágenes monocromáticas se las denomina *imágenes en escala de grises*.

Las imágenes de radiología convencional y de TC son imágenes en escala de grises, donde cada valor de píxel conserva un valor correspondiente a la tonalidad de gris. Las imágenes establecidas de este formato están compuestas de sombras de grises, partiendo desde el negro más profundo cambiando gradualmente hasta el color blanco. Es decir, son imágenes monocromáticas de 256 colores, $z \in [0, 255]$ como se puede apreciar en la Figura 31. Como hemos visto una imagen en escala de grises se puede representar como un conjunto de tonalidades de gris pero también se les puede dar color a ese conjunto de tonalidades. El uso de color en el procesamiento de imágenes es de mucha utilidad, como por ejemplo, para representar una imagen en un mapa de colores. En análisis de imágenes, el **color** es un eficaz descriptor, simplifica

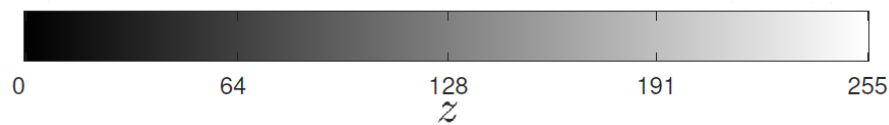


Figura 31: Escala de grises, $z \in [0, 255]$.

la identificación y extracción de detalles de una imagen, y por otro lado, el ojo humano puede distinguir una amplia gama de colores comparado con la escala de gris [1]. Esto lo podemos conseguir mediante una **paleta de colores**.

4.4. Paleta de color

Al hablar de color es conveniente mencionar tres propiedades que lo definen y son: *tono*, *saturación* y *brillo*. Tono es el atributo que diferencia un color de otro. A los diferentes tonos o matices se los puede representar en un círculo cromático. La saturación hace referencia a la **intensidad** o grado de pureza de cada color. Su valor máximo corresponde a un color puro o comúnmente se dice que es un color “intenso” o muy fuerte y cuando adquiere su mínimo valor corresponde a un tono de **gris** (Figura 32). Y finalmente el **brillo** o luminosidad, propiedad que será fundamental en el transcurso de esta investigación, indica la claridad u oscuridad de un color. Entre más brillo nos acercamos al blanco y menos al negro.

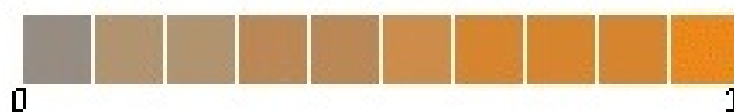


Figura 32: Saturación del color naranja, entre más nos acerquemos a 1 el color será más puro.

Estos atributos cambian y hacen exclusivo a cada color. Cualquier color con su máximo porcentaje de saturación (extremo derecho de la Figura 32) corresponderá a un color puro y poseerá su mayor pureza con un 100 % de brillo, y con el mínimo porcentaje de brillo será negro absoluto. En cambio, un color con un mínimo porcentaje de saturación (extremo izquierdo de la Figura 32) pertenecerá a un tono de gris y cambiará a un blanco absoluto con un 100 % de brillo y se convertirá en un negro absoluto con un 0 % de brillo.

En este estudio las imágenes monocromáticas (imágenes TC) son coloreadas por medio

de una **paleta de color** asignándoles un determinado color basándonos en las propiedades del nivel de intensidad de gris de la imagen original de TC. A esta técnica de conversión de escala de grises a color a menudo se lo conoce como *pseudocolor*. El uso principal del pseudocolor es para visualización e interpretación humana, ya que los humanos pueden discernir miles de tonos e intensidades de color [10]. En ROOT se puede acceder mediante la función `gStyle->SetPalette(número)`, donde *número* corresponde a una determinada paleta que se desea hacer uso de las 63 paletas de alta calidad disponibles, cada una de ellas predefinidas con 256 colores. Esta paleta hace referencia a la propiedad de brillo o luminosidad, es decir, la cantidad de luz que refleja una superficie.

La escala de grises también se puede medir por porcentajes, desde 0 % color negro a 100 % igual a blanco. Por fines prácticos y de simplificación de cifras se ha ajustado la escala que va desde 0 para el color negro a 1 que corresponde el color blanco (en lugar de porcentajes o de 0 a 255 como lo denota la expresión 17). En esta investigación se usó la paleta número 53, un tono que lleva por nombre “KDarkBodyRadiator”, en otras palabras, se ocupó la paleta `gStyle->SetPalette(53)` que la podemos ver en la Figura 33. Al convertir a *modo color indexado* el programa almacena y genera los índices de los colores simulando los colores de la imagen original con los colores disponibles de la paleta. Debido a esto se observa un notorio cambio en la gama de colores en el mapa de calor, ya que se consigue la misma luminosidad pero con otra paleta. De este modo la imagen adquiere colores cálidos. La finalidad de este modelo de color es facilitar la especificación de colores (donde cada color se representa por un punto conocido).

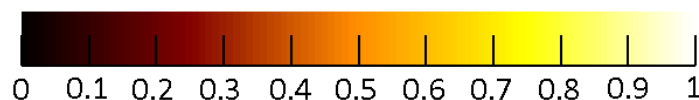


Figura 33: Paleta de colores usada en esta investigación con un rango de [0,1]

4.5. Imágenes de TC de tórax, mapas de calor e histogramas

4.5.1. Imágenes de TC de tórax para el diagnóstico de COVID-19

Como ya lo hemos afirmado la TC es una herramienta de diagnóstico por imágenes que acopla los rayos-X con la tecnología computarizada para obtener una imagen transversal en

escala de grises muy detallada de alguna área específica del cuerpo que se desee radiografiar. En una imagen de TC el color negro intenso muestra que los rayos no han sido atenuados mientras que el color blanco señala que la densidad alta del tejido no ha permitido pasar la radiación.

Con la finalidad de poder interpretar las imágenes citaremos 4 densidades radiológicas básicas que se pueden encontrar en el cuerpo humano, estas se producen por los fenómenos de absorción, dispersión y transmisión. En nuestro caso, las diferentes densidades relativas del tórax se identifican como tonos que van desde el negro, pasando por diferentes tonalidades de grises y que llegan al blanco. De menor a mayor densidad tenemos: (1) **Aire o gas**, (2) **Grasa o tejido adiposo**, (3) **Agua o tejidos blandos** y (4) **Hueso o calcio**. Ver Figura 34.



Figura 34: Cuatro densidades básicas de una TC

1. La densidad del aire se lo visualiza de color **negro**, decimos que es **radiolúcido**. Esto lo podemos ver en los pulmones, bronquios y los alvéolos pulmonares [4].
2. El tejido adiposo conocido popularmente como grasa corporal se observa de color **gris oscuro**.
3. Se considera la densidad de agua ya que órganos como el corazón, diafragma, vasos sanguíneos tienen una densidad similar que el agua [4]. Se los observa en las TC de un tono blanquesino o **gris claro**.

4. La densidad del hueso presenta un color **blanco**, en otras palabras, es completamente **radiopaca**.

La TC de tórax se la interpreta como si el paciente estuviera en frente de nosotros, es decir, el lado derecho de la imagen corresponde al lado izquierdo del paciente y viceversa [4].

4.5.2. Mapas de Calor

Decimos que un mapa de calor es el “mejoramiento” y manipulación de una imagen de TC para que sea más favorable y útil para un posterior análisis. Entonces, como punto de partida se analizaron las intensidades de varias imágenes que se las tomó como muestra, a estas imágenes las llamamos “mapas de calor”, en estos mapas están representados los niveles de **intensidad** de cada píxel con n bits, que para nuestro caso es $n=8$ (para obtener estos mapas se utilizó un programa en C++ usando ROOT).

Como hipótesis podemos argumentar que las partes con una intensidad mayor son los hallazgos (daños) más comunes citados anteriormente que tienen los pulmones asociados con la enfermedad del COVID-19. Y asociado a estas imágenes hemos elaborado los respectivos histogramas que muestran una distribución de área de la intensidad. En la Figura 35 se muestra la TC original junto con el mapa de calor que se obtuvo con el programa ROOT.

Un mapa de calor (Figura 35b) es una representación gráfica de datos en la que los valores individuales contenidos en una matriz se representan como colores. Los puntos de datos están definidos por la intersección de los ejes x e y y un tercer valor z que determina el color del punto de datos [11]. Como lo hemos afirmado anteriormente los píxeles de la imagen digital están dispuestos en forma de rejilla (adjuntar imagen de la página 37) dicho de otra manera, una imagen de TC es discreta, no existen valores de intensidad entre los valores discretos de las coordenadas x e y que confluyan con las intersecciones tanto horizontales y verticales de los valores discretos [1].

Recordemos que un punto puede ser representado por un par ordenado, por lo tanto, en los mapas de calor la distribución espacial de los píxeles de la imagen que hemos transformado se les ha designado arbitrariamente en el rango de $[-1, 1]$ en ambos ejes, con esto nos aseguramos

de conocer el valor de intensidad de un píxel en una posición cualquiera o en cualquier punto intermedio entre píxeles (podríamos decir que son coordenadas pseudo discretas). Mediante esta transformación es imprescindible obtener los valores de intensidad de manera que la imagen original aparezca asociada con los valores de intensidad extraídos de la imagen de TC original.

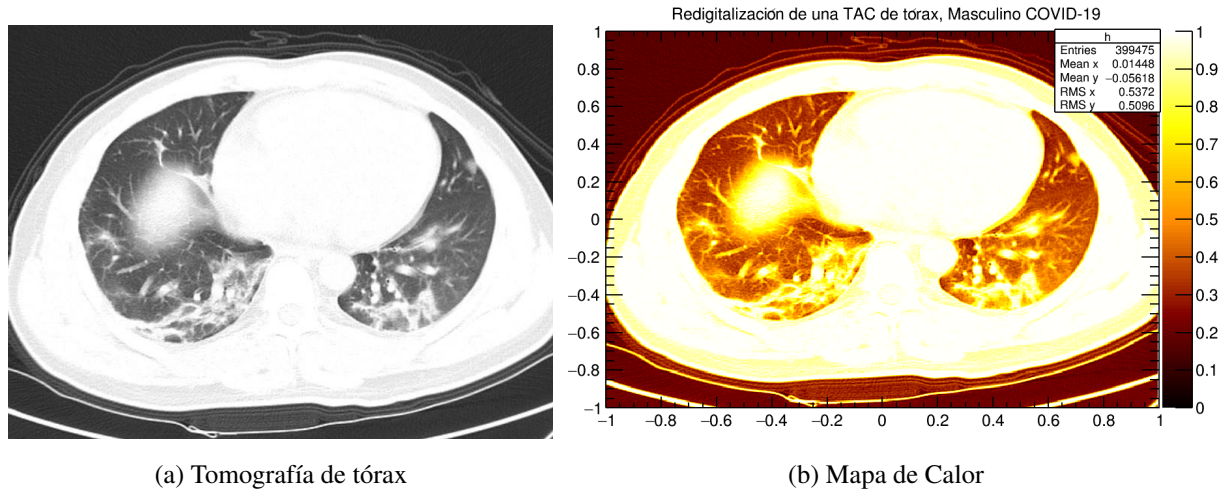


Figura 35: Transformación de una TC original a un mapa de calor

4.5.3. Intensidades de las imágenes como variables aleatorias

En esta investigación tratamos a las intensidades de las imágenes como cantidades aleatorias. Dado el nivel de intensidad z_k , donde $k = 0, 1, 2, \dots, L - 1$ denotan los valores de todas las posibles intensidades en una imagen digital $M \times N$ [10]. La **probabilidad**, $p(z_k)$ de que el nivel de intensidad z_k aparezca en la imagen se denota como

$$p(z_k) = \frac{n_k}{MN} \quad (20)$$

donde n_k es el número de veces que la intensidad z_k ocurre en la imagen y MN es el número total de píxeles en la imagen. Evidentemente, la suma de $p(z_k)$ para todos los valores de k es 1. Los componentes de $p(z_k)$ son estimaciones de la probabilidad de ocurrencia de los niveles de intensidad en una imagen.

$$\sum_{k=0}^{L-1} p(z_k) = 1 \quad (21)$$

4.5.4. Histogramas de Intensidad

El histograma de una imagen es una función discreta (función donde la variable independiente solo puede tomar determinados valores) y se lo representa por medio de un diagrama de barras cuyas abscisas denotan los niveles de gris de una imagen (z_k), y las ordenadas, las frecuencias relativas de los diferentes niveles de gris $p(z_k)$ [12], dicho de otro modo, la cantidad de entradas o píxeles $h(z_k)$ asociados a cada nivel de gris dividido por el número de píxeles total de la imagen ($M \times N$).

Entonces, dado el nivel de intensidad z_k , donde $k = 0, 1, 2, \dots, L - 1$ denotan las intensidades de una imagen digital $f(x, y)$. El histograma no-normalizado de f se define como

$$h(z_k) = n_k \quad \text{para } k = 0, 1, 2, \dots, L - 1 \quad (22)$$

donde n_k es el número de píxeles en f con intensidad z_k y a las subdivisiones de la escala de intensidad se les denomina *histogram bins* o simplemente *bins*. Similarmente, el **histograma normalizado** de f o la distribución de probabilidad $p(z_k)$ del k -ésimo valor de intensidad viene dada por la siguiente expresión

$$p(z_k) = \frac{h(z_k)}{MN} = \frac{n_k}{MN} \quad (23)$$

donde $h(z_k)$ es el número de veces que la intensidad z_k ocurre en la imagen. En la presente tesis, trabajaremos con histogramas normalizados a los que simplemente nos referimos como histogramas. Por consiguiente, una imagen tomográfica está conformada por diferentes valores en un mapa de intensidad que se pueden analizar estadísticamente al representarlos gráficamente como un histograma. Por ejemplo, observamos en una imagen de TC oscura (Figura 36a) que la mayor población de los bins del histograma están más concentrados en el extremo inferior (izquierda) de la escala de intensidad

De igual manera, la mayor población de bins en una imagen de TC clara (Figura 37) esta sesgado al extremo superior (derecha) de la escala de intensidad.

Para una imagen de 8 bytes, siendo el rango de los niveles de tonos de gris será $k \in [0, 256 - 1]$, se cumple que la suma de las frecuencias $h(z_k)$ da como resultado el número de píxeles de la imagen ($M \times N$), matemáticamente lo podemos expresar como

$$\sum_{k=0}^{255} h(z_k) = MN \quad (24)$$

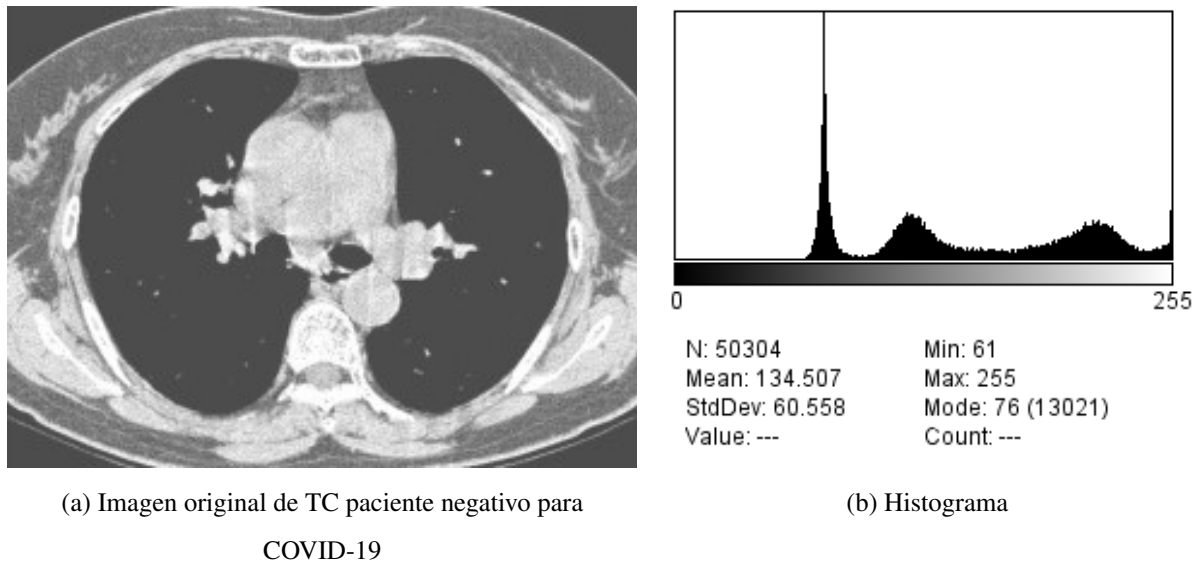


Figura 36: Imagen de TC oscura y su correspondiente histograma. El eje horizontal del histograma son valores de z_k y el vertical son valores de $p(z_k)$.

La ecuación 21 quedaría así

$$\sum_{k=0}^{255} p(z_k) = 1 \quad (25)$$

Como lo describimos anteriormente, el eje horizontal de los histogramas representan los niveles de gris z_k que en la presente investigación por fines prácticos y de simplificación lo hemos normalizado a $k \in [0, 1]$ y en el eje vertical están representados los valores de $p(z_k)$, esto es el número de píxeles o de entradas correspondientes a un nivel de intensidad. Se aprecia en la Figura 38c el histograma de intensidad de la imagen de TC de la Figura 38a con los niveles de intensidad variando en el rango de valores de 0 a 1 y su correspondiente mapa de calor.

Basándonos en los mapas de calor, que contienen la referencia entre -1 y 1 para ambos ejes (Figura 38b), se puede calcular la intensidad de una región en específico de la imagen de TC. Por ejemplo, si deseamos calcular la intensidad de la esquina superior derecha de la imagen se debe buscar las coordenadas, el eje x en el programa representa X_1 así $0.8 < X_1 < 1$ y para el y representa X_2 así $0.85 < X_2 < 1$.

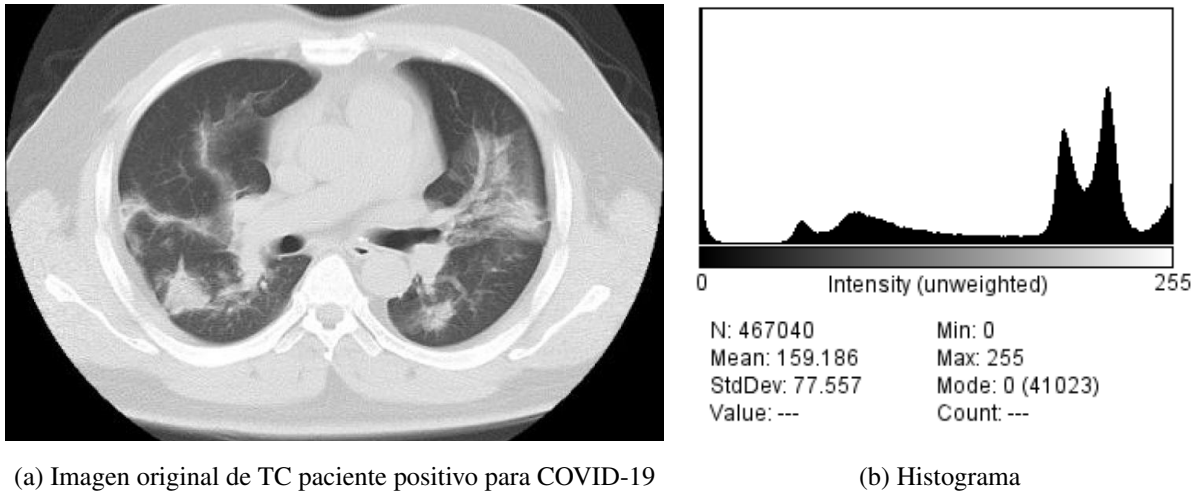


Figura 37: Imagen de TC clara y su correspondiente histograma. El eje horizontal del histograma son valores de z_k y el vertical son valores de $p(z_k)$.

4.5.5. Propiedades estadísticas de un histograma: Media y desviación estándar

Una propiedad que nos informa sobre la distribución de los niveles de gris de una imagen basándonos en su respectivo histograma y que representa el análisis principal de esta investigación es la **media** de las intensidades ($\mu_{\bar{z}}$). La media que es el valor medio de los niveles de gris nos brinda información sobre el **brillo** general de la imagen de TC, viene definida por la siguiente expresión

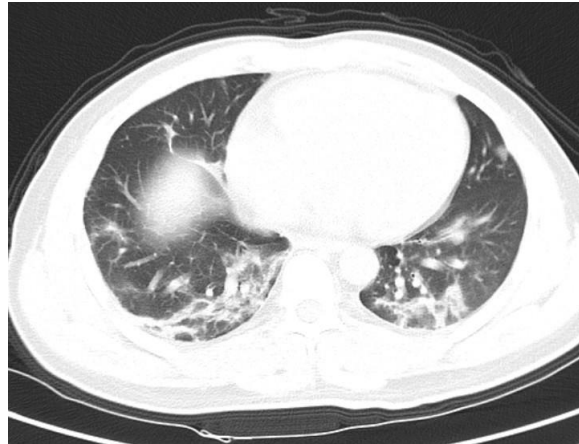
$$\mu_{\bar{z}} = \sum_{k=0}^{L-1} \frac{z_k h(z_k)}{MN}, \quad (26)$$

insertando la ecuación 23 en la anterior expresión obtenemos la media en función de la probabilidad

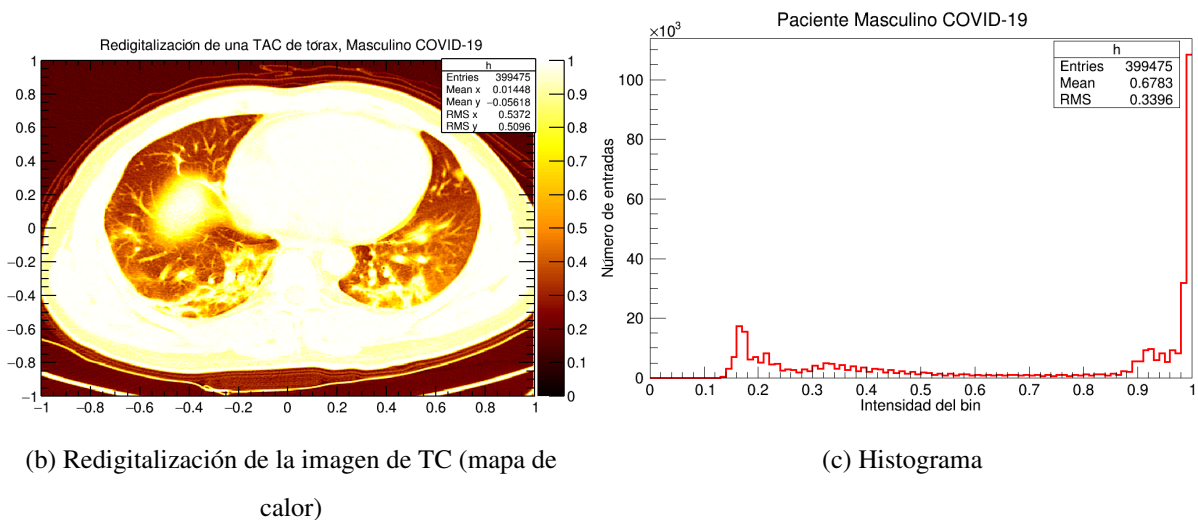
$$\mu_{\bar{z}} = \sum_{k=0}^{L-1} z_k p(z_k) \quad (27)$$

Una imagen TC brillante adquirirá un valor de media alto y por el contrario una imagen de TC oscura tendrá una media baja. Esta diferencia de valores lo podemos apreciar claramente en los histogramas de ruido de la Figura 41. Similarmente la varianza de las intensidades es

$$\sigma^2 = \sum_{k=0}^{L-1} (z_k - \mu_{\bar{z}})^2 p(z_k) \quad (28)$$



(a) Imagen de TC paciente positivo para COVID-19



(b) Redigitalización de la imagen de TC (mapa de calor)

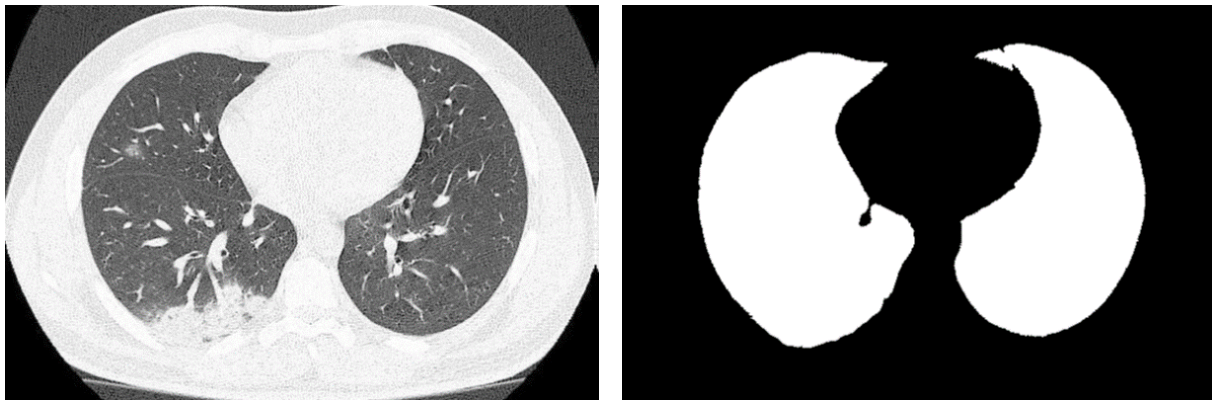
(c) Histograma

Figura 38: Imagen original con su respectiva redigitalización e histograma.

4.6. Segmentación de imágenes

El algoritmo de segmentación con el que se trabajará se basa en dos propiedades básicas de los valores de intensidad de una imagen de TC: **discontinuidad** y **similitud**. La primera se refiere a dividir las imágenes en regiones en función de los cambios bruscos de intensidad como los bordes (aquí separamos en ruido claro, ruido oscuro y región de interés). Los límites de las regiones son lo suficientemente diferentes entre sí y del fondo o ruido lo que permite la detección de límites basada en las intensidades en intensidad. Y la segunda como su nombre lo indica se basa en dividir una imagen en regiones que son similares de acuerdo a criterios que los iremos definiendo.

Por ejemplo, en la Figura 39a muestra una imagen típica de TC. Si nos centramos estrictamente en las regiones dentro del cuerpo humano (excluyendo partes ajenas al objeto de estudio) se observa una **discontinuidad de intensidades** entre el tejido adiposo (grasa corporal) y los lóbulos pulmonares. Para segmentar la imagen designamos un nivel (blanco para los píxeles que están en el límite e interior de los lóbulos pulmonares) y otro nivel (negro) para a todos los puntos fuera del límite como se lo muestra en la Figura 39b. En otras palabras, si un píxel está sobre el límite o dentro aplicamos la condición de etiquetarlo como blanco en otro caso se lo etiqueta como negro. Observamos que esta condición es **verdadera** para todos los puntos etiquetados como blanco y negro en la Figura 39b.



(a) Imagen de TC

(b) Resultado de la Segmentación

Figura 39: Imagen de TC con el resultado de la segmentación

El enunciado anterior, lo podemos expresar matemáticamente de la siguiente manera: Sea R la región espacial entera ocupada por una imagen de TC. Se podría decir que la segmentación es un proceso que divide a R en n subregiones $R_1, R_2, R_3, \dots, R_n$, tal que

$$Q(R_i) = \text{Verdadero} \quad \text{para } i = 1, 2, 3, \dots, n \quad (29)$$

condición que se cumple si todos los píxeles en R_i tienen la misma intensidad.

El algoritmo recibe como entrada la imagen de TC en escala de grises, luego a esta imagen se le designa un pseudocolor. Después seleccionamos un número previamente calculado en el rango de $[0, 255]$ que lo denominaremos **umbral** [13]. A este valor se lo compara con el nivel de intensidad de cada píxel con la que está formada la imagen, es decir en la posición $f(x, y)$



se comprueba si el valor es menor o igual que el umbral. En la matriz máscara, que tiene las mismas dimensiones que la imagen de entrada, se le asignan los valores de **Verdadero (V)** y **Falso (F)** establecido por la condición de umbralización. A continuación, se procede con la operación **AND** entre la imagen de entrada y la máscara. En las posiciones de la matriz donde se encuentre un valor **V** el valor de la imagen de entrada no se modifica. En cambio, si el valor es **F** el valor del píxel toma un valor de cero en la imagen de salida. La Tabla 4 y la Figura 40 muestran las operaciones que realiza el algoritmo para una matriz de 6x6 y umbral que esta en el rango de [0.45, 0.56].

Algoritmo	Función máscara/umbral
Entrada	Imagen TC tórax en escala de grises
Umbral	[0.45, 0.56]
Máscara	$0.45 \leq \text{Imagen de entrada} \leq 0.56$
Imagen de Salida	Imagen TC tórax AND Máscara

Tabla 4: Resumen del algoritmo para obtener las máscaras de las imágenes TC

Imagen de entrada	Umbralización	Imagen de Salida																																																																																																												
<table border="1" style="border-collapse: collapse; width: 100%;"> <tr><td>0.19</td><td>0.22</td><td>0.69</td><td>0.52</td><td>0.22</td><td>0.17</td></tr> <tr><td>0.24</td><td>0.31</td><td>0.46</td><td>0.52</td><td>0.72</td><td>0.19</td></tr> <tr><td>0.65</td><td>0.48</td><td>0.45</td><td>0.53</td><td>0.75</td><td>0.81</td></tr> <tr><td>0.67</td><td>0.54</td><td>0.50</td><td>0.47</td><td>0.49</td><td>0.79</td></tr> <tr><td>0.20</td><td>0.55</td><td>0.51</td><td>0.46</td><td>0.48</td><td>0.21</td></tr> <tr><td>0.18</td><td>0.29</td><td>0.47</td><td>0.52</td><td>0.56</td><td>0.23</td></tr> </table>	0.19	0.22	0.69	0.52	0.22	0.17	0.24	0.31	0.46	0.52	0.72	0.19	0.65	0.48	0.45	0.53	0.75	0.81	0.67	0.54	0.50	0.47	0.49	0.79	0.20	0.55	0.51	0.46	0.48	0.21	0.18	0.29	0.47	0.52	0.56	0.23	<table border="1" style="border-collapse: collapse; width: 100%;"> <tr><td>F</td><td>F</td><td>F</td><td>V</td><td>F</td><td>F</td></tr> <tr><td>F</td><td>F</td><td>V</td><td>V</td><td>F</td><td>F</td></tr> <tr><td>F</td><td>V</td><td>V</td><td>V</td><td>F</td><td>F</td></tr> <tr><td>F</td><td>V</td><td>V</td><td>V</td><td>V</td><td>F</td></tr> <tr><td>F</td><td>V</td><td>V</td><td>V</td><td>V</td><td>F</td></tr> <tr><td>F</td><td>F</td><td>V</td><td>V</td><td>V</td><td>F</td></tr> </table>	F	F	F	V	F	F	F	F	V	V	F	F	F	V	V	V	F	F	F	V	V	V	V	F	F	V	V	V	V	F	F	F	V	V	V	F	<table border="1" style="border-collapse: collapse; width: 100%;"> <tr><td>0</td><td>0</td><td>0</td><td>0.52</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0.46</td><td>0.52</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0.48</td><td>0.45</td><td>0.53</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0.54</td><td>0.50</td><td>0.47</td><td>0.49</td><td>0</td></tr> <tr><td>0</td><td>0.55</td><td>0.51</td><td>0.46</td><td>0.48</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0.47</td><td>0.52</td><td>0.56</td><td>0</td></tr> </table>	0	0	0	0.52	0	0	0	0	0.46	0.52	0	0	0	0.48	0.45	0.53	0	0	0	0.54	0.50	0.47	0.49	0	0	0.55	0.51	0.46	0.48	0	0	0	0.47	0.52	0.56	0
0.19	0.22	0.69	0.52	0.22	0.17																																																																																																									
0.24	0.31	0.46	0.52	0.72	0.19																																																																																																									
0.65	0.48	0.45	0.53	0.75	0.81																																																																																																									
0.67	0.54	0.50	0.47	0.49	0.79																																																																																																									
0.20	0.55	0.51	0.46	0.48	0.21																																																																																																									
0.18	0.29	0.47	0.52	0.56	0.23																																																																																																									
F	F	F	V	F	F																																																																																																									
F	F	V	V	F	F																																																																																																									
F	V	V	V	F	F																																																																																																									
F	V	V	V	V	F																																																																																																									
F	V	V	V	V	F																																																																																																									
F	F	V	V	V	F																																																																																																									
0	0	0	0.52	0	0																																																																																																									
0	0	0.46	0.52	0	0																																																																																																									
0	0.48	0.45	0.53	0	0																																																																																																									
0	0.54	0.50	0.47	0.49	0																																																																																																									
0	0.55	0.51	0.46	0.48	0																																																																																																									
0	0	0.47	0.52	0.56	0																																																																																																									
	AND	=																																																																																																												

Figura 40: Función máscara. Para este ejemplo el umbral está en el intervalo [0.45, 0.56]

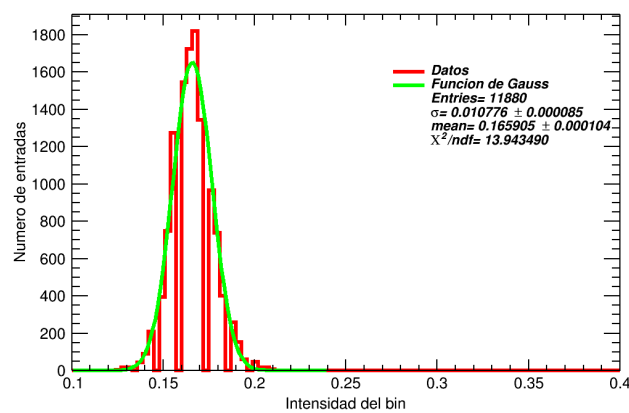
4.7. Histogramas de Ruido

El **ruido** de una imagen es la variación aleatoria de intensidad o color [14]. En el proceso de adquisición de imágenes es muy común que vengan acompañadas de ruido o que la calidad

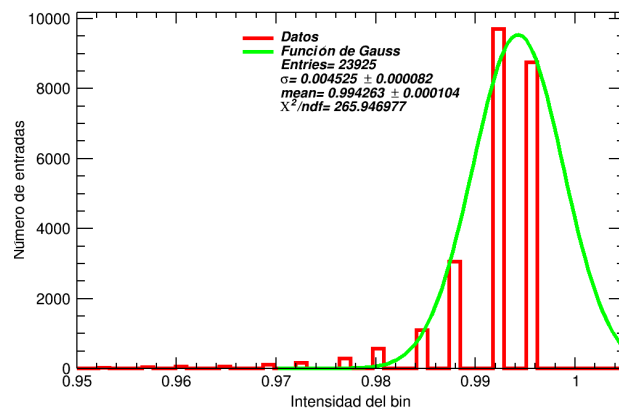


no sea la deseada lo que ocasiona inconvenientes para interpretar minúsculos detalles y características morfológicas de baja intensidad indispensables para un diagnóstico clínico óptimo y la medición del sesgo de un histograma, se decir, si la mayor parte de los valores de la distribución se encuentran ubicados a la izquierda o derecha del valor medio nos suministra información muy importante como lo veremos a continuación.

Como podemos observar en la Figura 41a los píxeles oscuros son los dominantes y en la Figura 41b se aprecia que los píxeles más claros son los dominantes con los niveles de intensidad están muy cercanos a 1.



(a) Ruido oscuro



(b) Ruido claro

Figura 41: Histogramas de Ruido

Entonces, las regiones oscuras de las imágenes de TC de tórax o las imágenes en general oscuras tendrán un histograma con agrupaciones o barras de intensidad primordialmente al ex-



tremo izquierdo de la gráfica. En cambio el histograma claro favorecerá a agruparse al extremo derecho de la gráfica (Ver Figura 41). En TC, durante el proceso de adquisición de imágenes si se disminuye la radiación ionizante aumenta la cantidad de ruido.

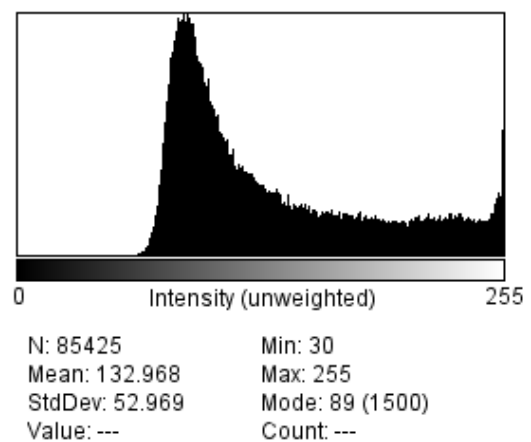
4.8. Región de Interés (ROI)

Las propiedades esenciales de los píxeles en las que nos centraremos son su posición y su nivel de intensidad si esta en la escala de grises o niveles de intensidad si analizamos los mapas de calor. En las imágenes TC de tórax se aprecia claramente ciertas áreas o agrupaciones de píxeles relacionados entre sí ya que presentan propiedades o características comunes, por ejemplo, el color. Llamaremos a estas agrupaciones **regiones**. Y cuando se analizan TC en general, se requiere investigar específicamente un área dentro de dicha imagen médica, a la que denominaremos **región de interés** (ROI, del inglés *region of interest*).

Específicamente la ROI de nuestras imágenes de TC son los tres lóbulos (secciones) con los que esta conformado el pulmón derecho y los dos lóbulos pulmonares en los que se divide el pulmón izquierdo. El histograma de las ROI esta representado como un “pico de montaña” (Figura 42b), es decir, el histograma no se corta ni tiende a sus extremos.



(a) Imagen original de TC paciente positivo para COVID-19



(b) Histograma de la región de interés

Figura 42: a) Imagen de TC, b) Histograma de la ROI

En la Figura 43 se realiza una comparación de las ROI (histogramas de color verde) de los ruidos para 50 imágenes de pacientes No COVID-19.

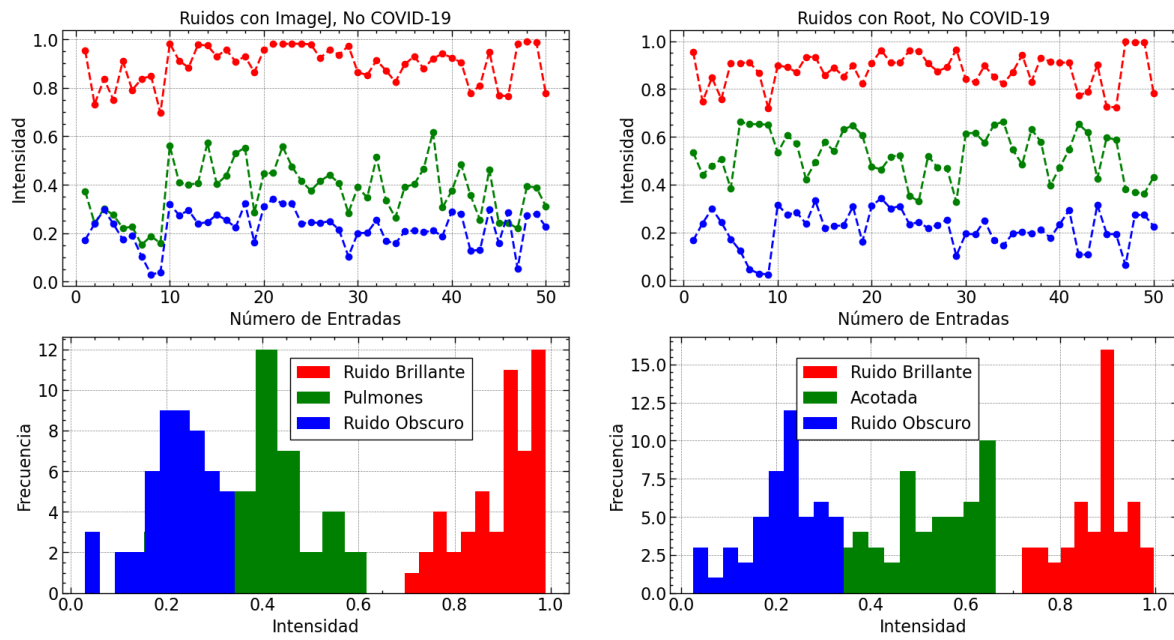


Figura 43: Comparación de las ROI (de color verde) con el ruido oscuro y claro usando ROOT e ImageJ



Capítulo 5

5. Resultados

5.1. Resultados Imágenes pacientes COVID-19

Como lo hemos indicado anteriormente, la interpretación de una tomografía torácica es un trabajo realizado por radiólogos con sólidos conocimientos en anatomía de tórax. En la presente investigación, basándonos en el poder de penetración de los rayos-X y en los fenómenos físicos que esta interacción produce, se ha reconocido una imagen de TC de tórax de pacientes sanos (anatomía normal) y las imágenes de TC de pacientes positivo para COVID-19 (alteraciones potencialmente graves). Se hizo el análisis con 133 pacientes diagnosticados de **positivo** para COVID-19. Estos pacientes se clasificaron en tres grupos: grupo no crítico (n=47) a los pacientes que se recuperaron de la COVID-19 y fueron dados de alta, grupo crítico (n=45) pacientes ingresados a la unidad de cuidados intensivos (UCI) y grupo de muerte (n=41) a las personas que fallecieron. Para los pacientes sanos o que no padecen la enfermedad de coronavirus se trabajó con una muestra de 70 pacientes.

Características	Gravedad de la enfermedad		
	No Crítico (Asintomático) (n=47, 35.3 %)	Crítico (n=45, 33.8 %)	Defunción (n=41, 30.8 %)
Edad (años)	32 (21, 50)	43 (33, 54)	60 (50, 69)
Género, masculino	25 (53.2 %)	23 (51.1 %)	24 (58.5 %)
Síntomas iniciales			
Fiebre	9/22 (40.9 %)	32 (72.1 %)	33 (80.5 %)
Tos	13/22 (59.01 %)	26 (57.8 %)	30 (73.17 %)
Dificultad para respirar	0/22 (0 %)	2 (4.4 %)	7 (17.1 %)
Comorbilidad	5 (10.6 %)	10 (22.22 %)	17 (41.5 %)
UCI/VMI	0 (0 %)	2 (4.4 %)	16 (39.0 %)

Tabla 5: Características básicas de los 133 pacientes con COVID-19 estudiados.

Como se demuestra en los porcentajes de la Tabla 5 los hombres tienen una mortalidad mayor que las mujeres. Además, existe una proporción importante de pacientes asintomáticos, 25 de los 47 del grupo No Crítico no presentaron síntomas lo que corresponde al 53.2 % de ese grupo. Las siguientes enfermedades previas en los pacientes aumentan el riesgo de enfermedad grave, por ejemplo, enfermedad cardiovascular, diabetes mellitus, hipertensión arterial, tabaquismo, cáncer, enfermedad renal crónica y obesidad. Otro factor importante es la edad avanzada que se asocia con una mayor letalidad del virus.

Se determinó que los rasgos característicos de los pacientes infectados pueden aparecer en forma de un infiltrado localizado con opacidad o consolidación. Estos infiltrados pueden afectar a cualquier lóbulo pulmonar [4]. Las opacidades borrosas en vidrio deslustrado observadas en las TC son probablemente debidas al síndrome de dificultad respiratoria aguda (SDRA), siendo una de las principales causas de los pacientes críticos lo que los llevó a la unidad de cuidado intensivos. Otro signo característico fue el *patrón de empedrado*. El 74 % de los pacientes del grupo de muerte requirió ventilación mecánica invasiva, es decir, los pacientes que requirieron esta intervención presentaron un alto índice de mortalidad.

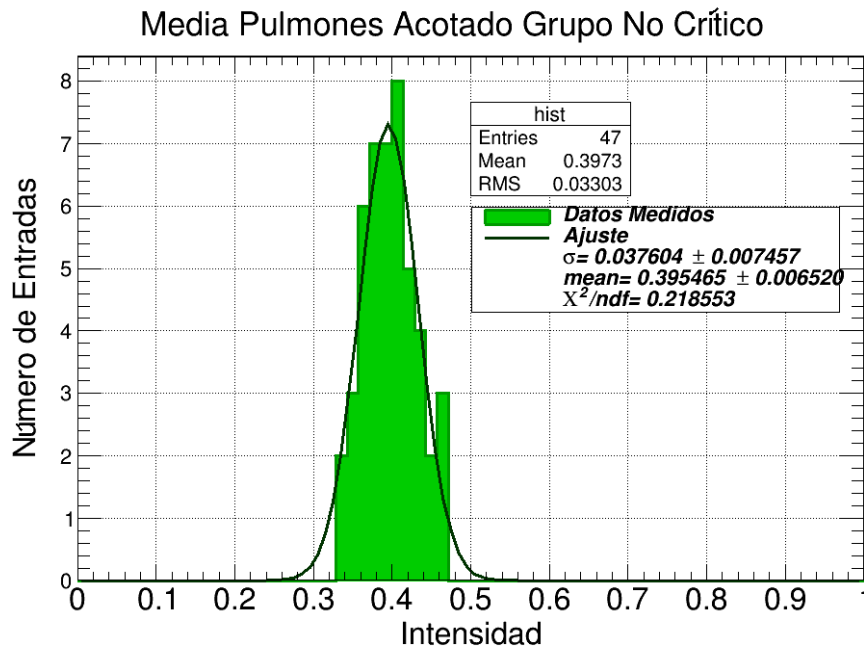


Figura 44: Histograma de la región de interés del **grupo no crítico** de una muestra de 47 pacientes.

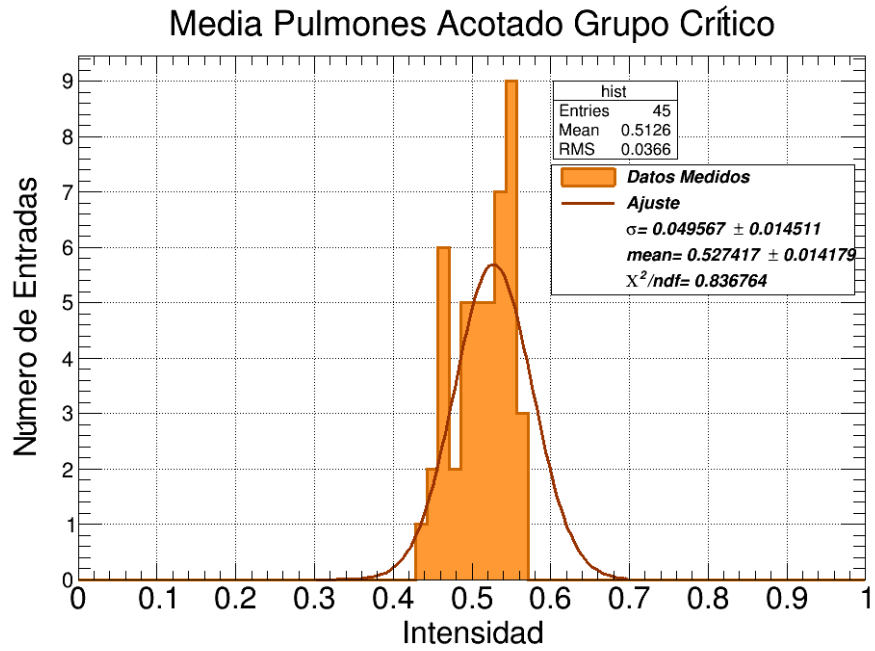


Figura 45: Histograma de la región de interés del **grupo crítico** de una muestra de 45 pacientes.

5.1.1. Resultado General para pacientes COVID-19

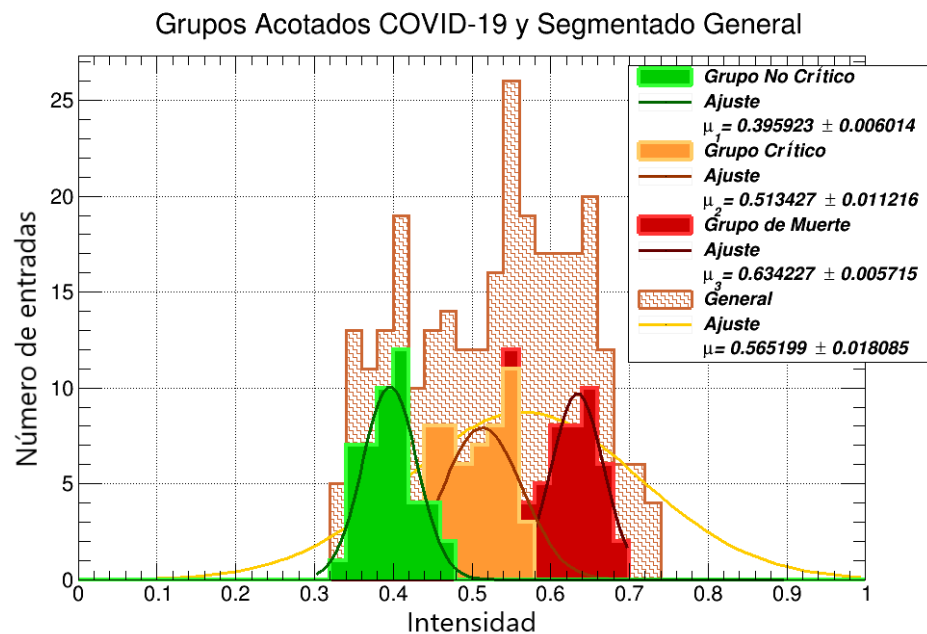


Figura 47: Histograma de la región de interés de los tres grupos estudiados de una muestra de 133 pacientes.

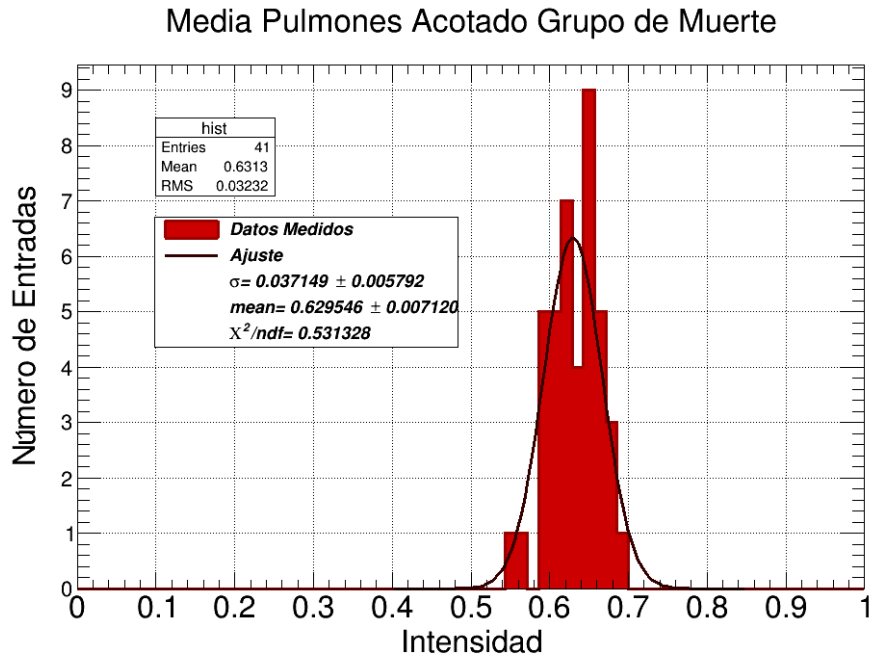


Figura 46: Histograma de la región de interés del **grupo de defunción** de una muestra de 41 pacientes.

Debido a su manejabilidad matemática y a su comprobado uso en procesamiento de imágenes se ha ocupado el modelo **gaussiano**. La función de densidad de probabilidad Gaussiana de una variable aleatoria z es la muy conocida ecuación

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-\mu)^2}{2\sigma^2}} \quad (30)$$

donde z representa la intensidad, μ es la media de los valores de z y σ es su desviación estándar.

Alguna de las propiedades de la curva normal es que tiene sus puntos de inflexión en $x = \mu \pm \sigma$. En las tablas 6 y 7 se muestran estos resultados.

Intensidad	Grupos COVID-19		
	No crítico	Crítico	Defunción
μ	0.3980	0.5130	0.6280
σ	0.0315	0.0343	0.0293

Tabla 6: Valores de intensidad de pacientes diagnosticados con COVID-19

Intensidad	Grupos No COVID-19		
	Corte 1	Corte 2	Corte 3
μ	0.3789	0.5148	0.6190
σ	0.0479	0.0316	0.0377

Tabla 7: Valores de intensidad de pacientes sin COVID-19

5.1.2. Intervalos de Confianza

Tanto a los pacientes diagnosticados de positivo para COVID-19 y a los pacientes sanos se los separó en tres grupos claramente definidos. Una vez obtenido la media de intensidad (μ) y la desviación estándar (σ) es necesario *estimar* entre que valores alrededor de la media se encuentra un rango superior y un rango inferior de cada grupo estudiado. Estos intervalos nos ayudarán a cuantificar la incertidumbre que se tiene alrededor del auténtico valor.

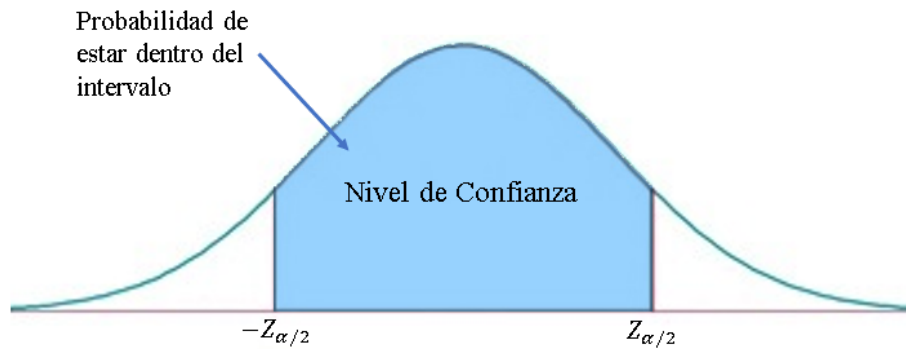


Figura 48: Intervalo de confianza

$$P(-Z_{\alpha/2} < Z < Z_{\alpha/2}) = 1 - \alpha \quad (31)$$

Donde

$$Z = \frac{\bar{X} - \mu}{\sigma} \quad (32)$$

Sustituyendo esta última expresión (32) en la ecuación (31) obtenemos

$$P\left(-Z_{\frac{\alpha}{2}} < \frac{\bar{X} - \mu}{\sigma} < Z_{\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$P(-Z_{\frac{\alpha}{2}} \sigma < \bar{X} - \mu < Z_{\frac{\alpha}{2}} \sigma) = 1 - \alpha$$

$$P(-\bar{X} - Z_{\frac{\alpha}{2}} \sigma < -\mu < -\bar{X} + Z_{\frac{\alpha}{2}} \sigma) = 1 - \alpha$$

$$P(\bar{X} + Z_{\frac{\alpha}{2}} \sigma > \mu > \bar{X} - Z_{\frac{\alpha}{2}} \sigma) = 1 - \alpha$$

reordenando nos queda finalmente

$$P(\bar{X} - Z_{\frac{\alpha}{2}} \sigma < \mu < \bar{X} + Z_{\frac{\alpha}{2}} \sigma) = 1 - \alpha \tag{33}$$

Nivel de confianza	α	$\alpha/2$	$Z_{\frac{\alpha}{2}}$	Grupos COVID-19		
				No Crítico	Crítico	Defunción
90 %	0.10	0.05	1.645	(0.346; 0.449)	(0.456; 0.569)	(0.579; 0.676)
95 %	0.05	0.025	1.960	(0.336; 0.459)	(0.445; 0.580)	(0.570; 0.685)

Tabla 8: Nivel de confianza de los grupos analizados

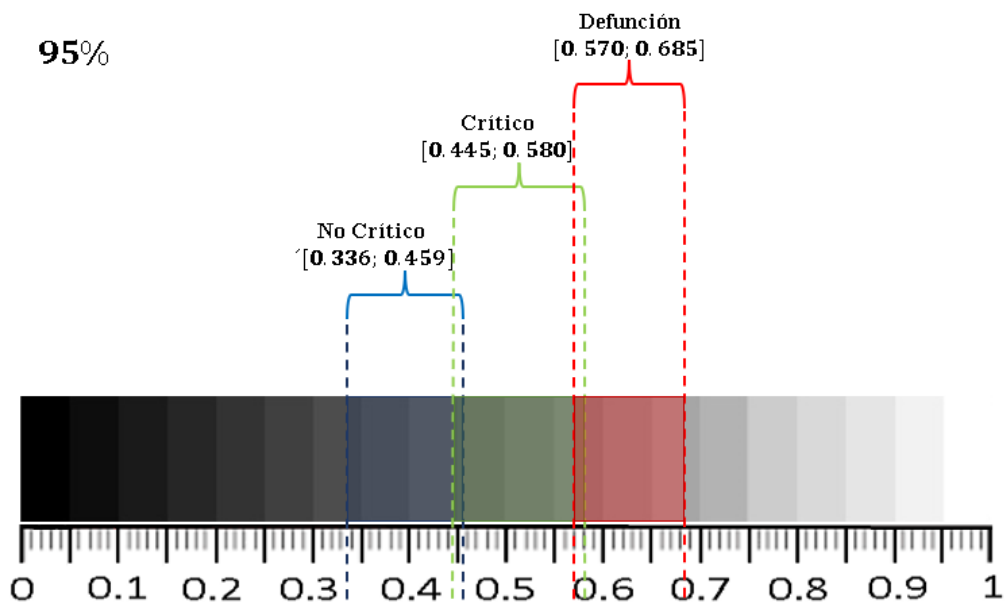


Figura 49: Esquema de intervalos de confianza de los grupos de COVID-19 para un nivel de confianza del 95 %

Para visualizar la dispersión de los datos se los ha representado en cuartiles en un diagrama de **caja y bigotes** (ver figura 51) donde se aprecian los límites de prolongación (bigotes). Estas cajas son muy útiles para comparar los tres grupos analizados.

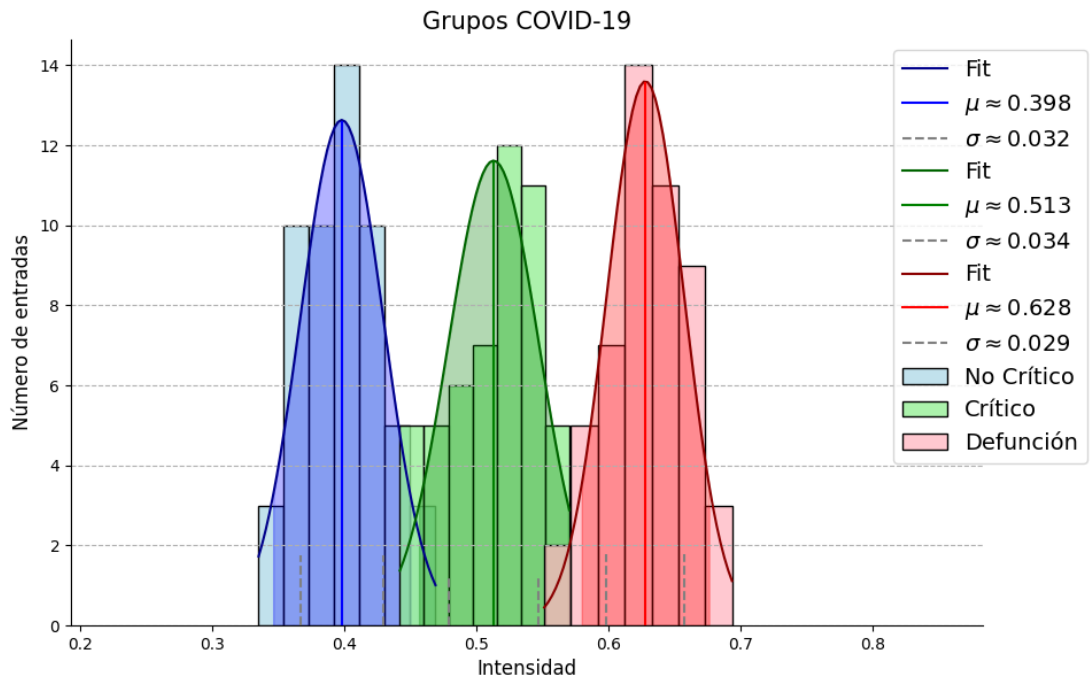


Figura 50: Se aprecia los intervalos de confianza de los grupos de COVID-19 para un nivel de confianza del 90 %

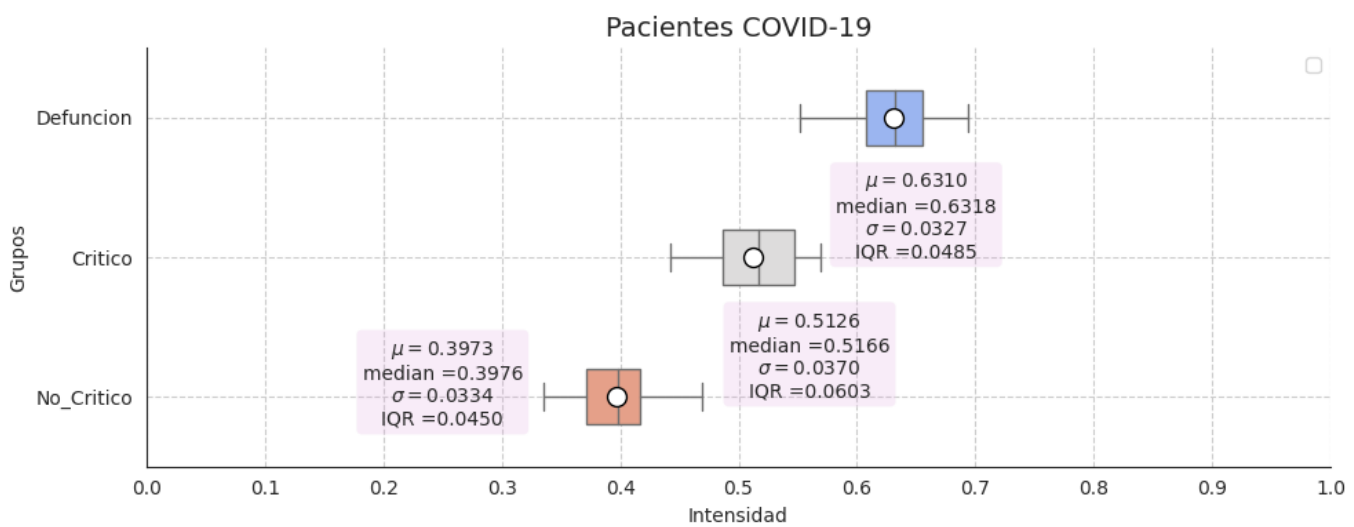


Figura 51: Diagrama de caja y bigotes donde se observa la dispersión de los datos.

5.2. Resultados Imágenes pacientes No COVID-19

Uno de los principales objetivos en el procesamiento de imágenes es eliminar la mayor cantidad posible de ruido. Una ventaja que posee el algoritmo utilizado es la elección de parámetros que ubica zonas fuera de los lóbulos para su posterior eliminación o minimización. La importancia de este proceso radica en el **filtrado** para obtener otra imagen con la mayor cantidad de píxeles pertenecientes a las regiones de interés. Los valores de lo que se ha considerado como ruido se muestran en la Tabla 9.

Intensidad	NO COVID-19		
	Oscuro	Acotado	Claro
μ	0.2110	0.5414	0.8742
σ	0.0376	0.049567	0.0371

Tabla 9: Valores de intensidad de los ruidos y de la región acotada de pacientes NO COVID-19

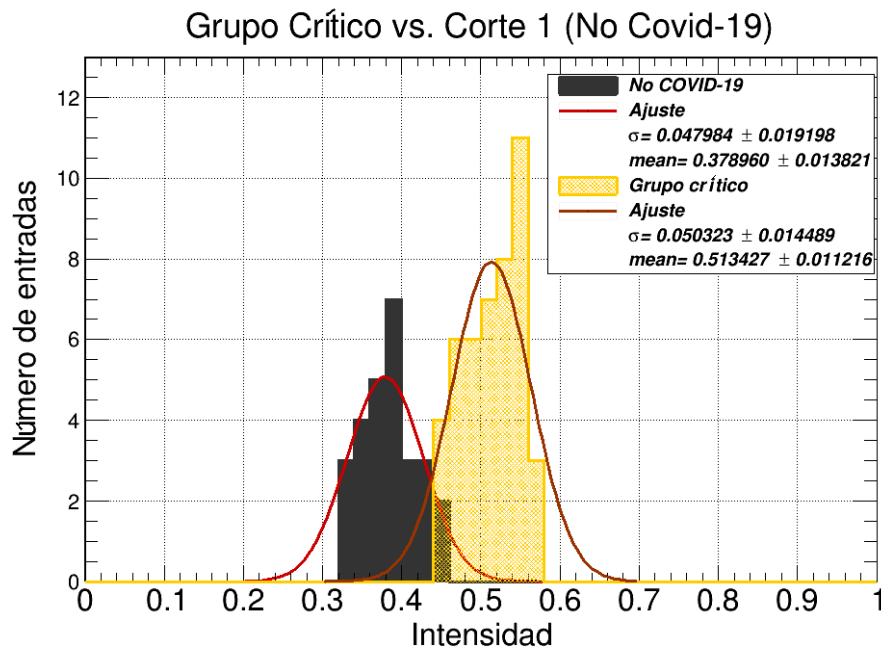


Figura 52: Histograma que compara los resultados del grupo crítico con el grupo de **corte 1**

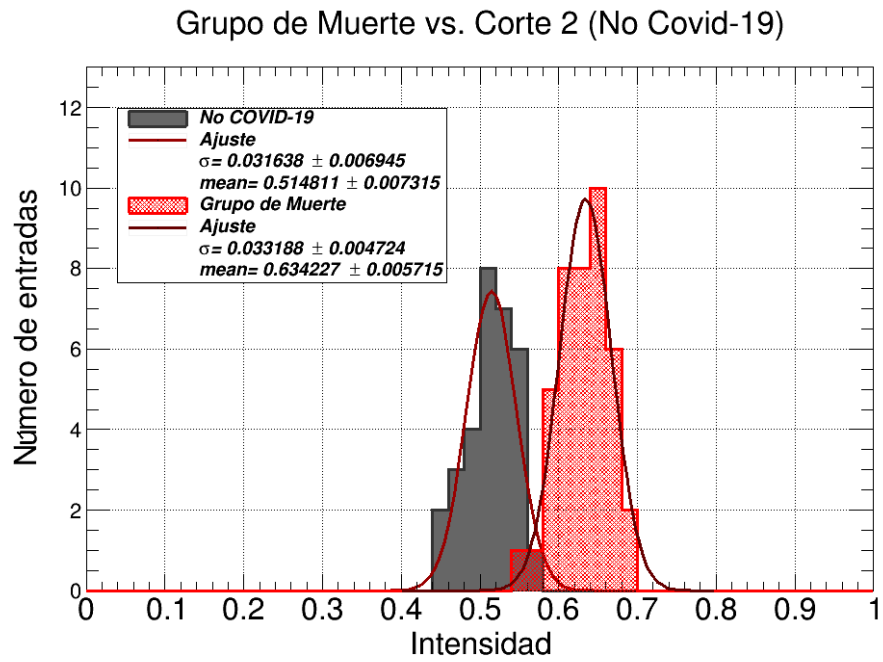


Figura 53: Histograma que compara los resultados del grupo de muerte con el grupo de **corte 2**

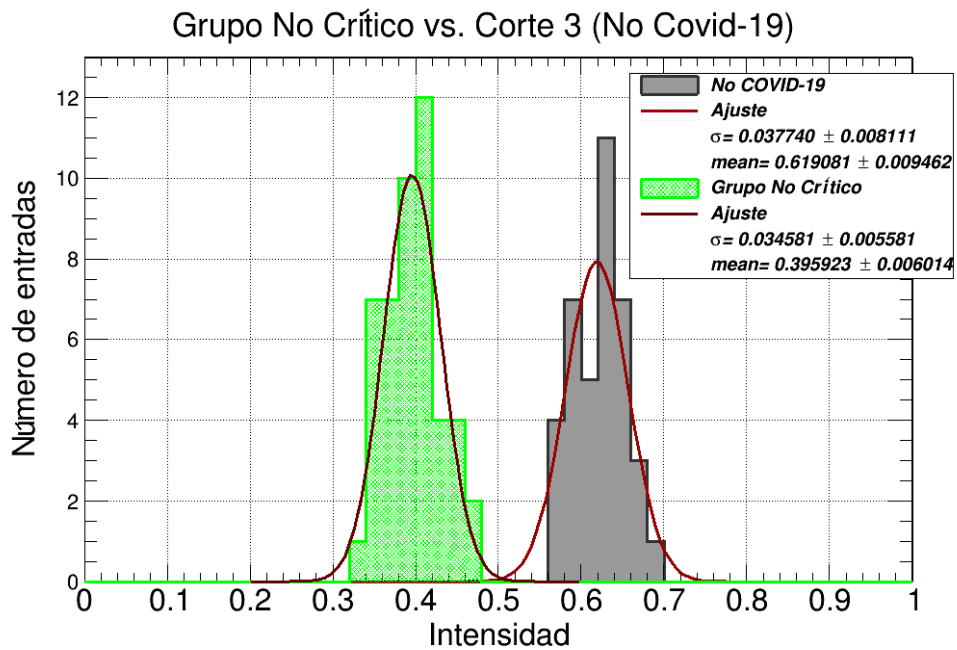


Figura 54: Histograma que compara los resultados del grupo no crítico con el grupo de **corte 3**



5.2.1. Resultado General para pacientes No COVID-19

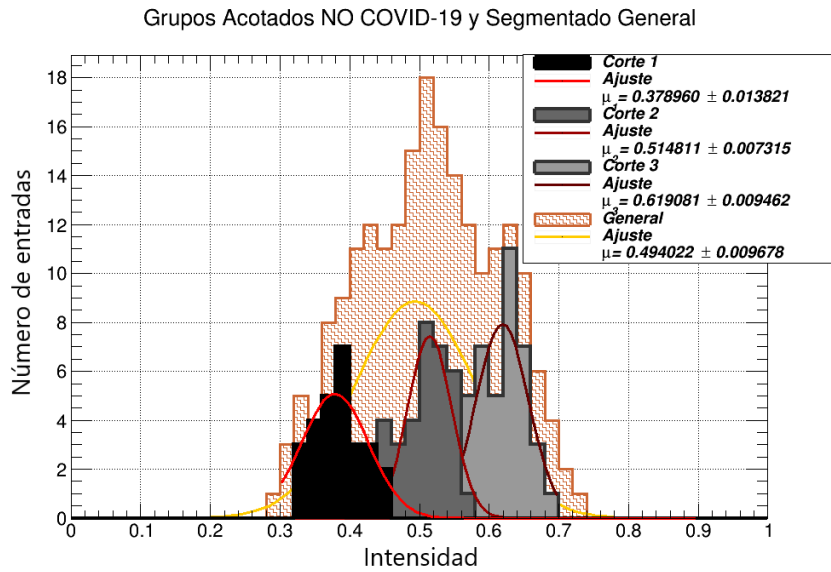


Figura 55: Histograma de la región de interés de los tres grupos estudiados de pacientes diagnosticados de negativo para COVID-19.

5.3. Comparación entre los tres grupos de COVID-19 vs. los tres grupos de NO COVID-19

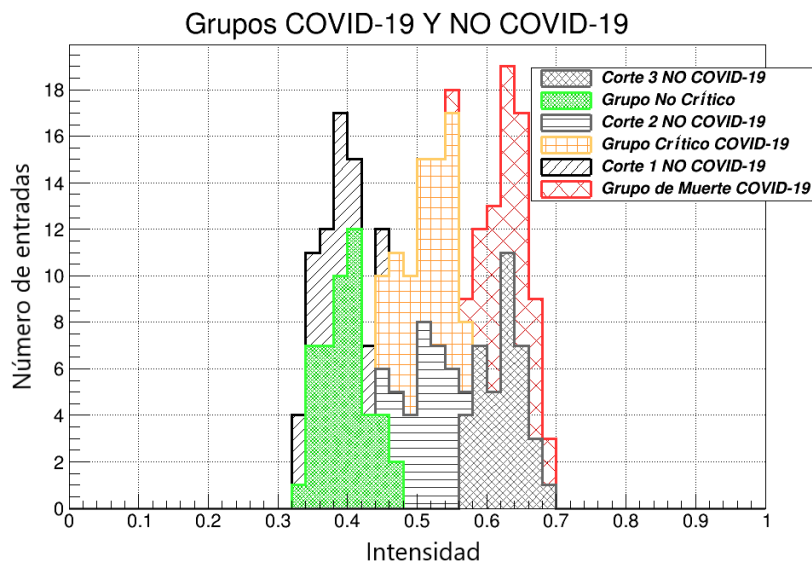


Figura 56: Histograma de la región de interés de los seis grupos analizados tanto de pacientes con COVID-19 como de pacientes sanos..



6. Análisis de Resultados

En este capítulo se presentan los valores obtenidos en el procesamiento de las imágenes. Además, se hace un análisis de los datos obtenidos y la discusión de los resultados.

Uno de los resultados más importantes fue distinguir a partir del mapa de intensidades el valor de dicha intensidad tanto de tejido pulmonar sano como de tejido pulmonar afectado por COVID-19, como se demuestran en los histogramas previos. Aunque los hallazgos histopatológicos requieren un análisis mucho más exhaustivo se ha logrado caracterizar a los pacientes sanos y los diagnosticados como positivos para esta enfermedad.

En los pacientes del grupo no crítico presentaron una afectación clínica leve que no requirió hospitalización y la mayoría pudo controlar la enfermedad desde casa. Presentaron manifestaciones clínicas secundarias con síntomas respiratorios leves. Como lo habíamos mencionado anteriormente la base de datos cuenta con un historial clínico de cada paciente, así que de acuerdo a los tres grupos estudiados de COVID-19 en la tabla 10 se enlistan los hallazgos radiológicos característicos que presentaron las personas infectadas con el virus.

Para las personas del grupo crítico y grupo de muerte si hubo la necesidad de hospitalización para un tratamiento y seguimiento de las complicaciones más comunes del COVID-19. El 65 % requirió atención en una unidad de cuidados intensivos (UCI) y para mejorar la oxigenación de estos pacientes se tuvo que recurrir a la ventilación mecánica invasiva. Los pacientes que precisaron ventilación mecánica invasiva manifestaron una alta mortalidad (entre el 60 y 80 %)

En los casos de los pacientes graves la hospitalización fue prolongada. Para estos dos grupos un factor determinante fue la edad avanzada (mayor a los 65 años), enfermedades cardiovasculares y diabetes. En el grupo crítico empieza la inflamación pulmonar y da paso a la evolución de la neumonía viral con tos, disnea (dificultad respiratoria) y un aumento progresivo de linfopenia (insuficiente cantidad de glóbulos blancos en la sangre). Para este grupo, desde los primeros síntomas hasta el ingreso hospitalario pasaron de 6 a 8 días.

Una de las características más comunes de las TC de los pacientes que días después fallecieron (grupo de muerte) fue que los dos pulmones tenían lesiones parcheadas dispersas con



opacidades en vidrio esmerilado (GGO) y engrosamiento difuso de las paredes alveolares. En algunos pacientes también del grupo de muerte se observó parches gris-blancos. Uno de los síntomas característicos de este grupo y la causa primordial de ingreso a las UCI fue el síndrome de dificultad respiratoria aguda (SDRA) cuyo patrón radiológico es el daño alveolar difuso (Parra Gordo et al., 2021).

Tipo de lesiones	Grupos		
	No crítico %	Crítico %	Defunción %
GGO	41.8	17.3	2.5
GGO más empedrado	15.4	14.8	9.4
Consolidación	4.6	10.6	4.6
GGO más consolidación	37.2	57.1	83.5

Tabla 10: Hallazgos radiológicos característicos de pacientes con COVID-19

Comentarios finales

- Si bien es recomendable analizar una secuencia de cortes (*slices*) de TC, a menudo un solo corte de TC contiene suficiente información clínica para tomar decisiones precisas.
- En la presente tesis se ha realizado un análisis de datos de intensidades de los píxeles. Se demostró que los pacientes diagnosticados para COVID-19 obtienen un media de intensidad de los pulmones mayor que los pacientes sanos.
- Desde la década de 1960 hasta la actualidad el procesamiento de imágenes ha avanzado a pasos agigantados y no para de crecer. Las computadoras han tenido una evolución sin precedentes, han introducido características que aumentan su utilidad, desempeño, unidades de almacenamiento más rápidas, etc. A pesar de que la experiencia adquirida en la interpretación de TC también ha progresado de forma considerable, la detección y caracterización de lesiones pulmonares sigue siendo un reto enorme.
- Este estudio presenta un método de tratamiento de datos dirigido a la identificación de áreas de mayor densidad en TC, con el fin de ayudar a detectar eficazmente las posibles



zonas de anormalidades en los pulmones partiendo de la caracterización cuantitativa de la segmentación de los lóbulos pulmonares.



Referencias

- [1] Krawczuk Maximiliano Martínez. «Segmentación de Imágenes Médicas». Facultad de Informática. Universidad Nacional de La Plata, feb. de 2007.
- [2] xingyi Yang et al. «COVID-CT-Dataset: A CT Image Dataset about COVID-19». En: *arXiv preprint arXiv:2003.13865v3* (2020).
- [3] Ashtatari S. et al. «Características tomográficas computarizadas de la neumonía por coronavirus-2019 (COVID-19) en tres grupos de pacientes iraníes: estudio de un solo centro». En: *Sociedad Española de Radiología Médica* RX.1285 (2020), pág. 10. DOI: 10.1016/j.rx.2021.03.002.
- [4] Pezzotti William. «Interpretación de la radiografía de tórax: algo más que blanco y negro». En: *Elsevier* (2014), págs. 30-37. DOI: 10.1016/j.nursi.2014.10.012.
- [5] Penelope Allys-Roberts y Jerry Williams. *Farr 's Physics for Medical Imaging*. Edinburgh: SAUNDERS Elsevier, 2008.
- [6] Paul Davidovits. *Physics in Biology and Medicine*. Fifth Edition. India: Academic Press:An imprint of Elsevier, 2019.
- [7] Nilesh Ugemuge, Yatish R. Parauha y S.J. Dhoble. «Synthesis and luminescence study of silicate based phosphors for energy saving light emitting diodes». En: *Energy Materials* (2021), págs. 445-480. DOI: 10.1016/b9780128237106.000170.
- [8] William Hendee R y Russell Ritenour. *Medical Imaging Physics*. 4.^a ed. New York: Wiley-Liss, 2002.
- [9] Agulló García José Antonio. «Algoritmos paralelos para la eliminación de ruido en imágenes médicas». Departamento de Ciencias de la Computación. Universidad de Alicante, jun. de 2019.
- [10] Richard E. Gonzalez. Rafael C. Woods. *Digital Image Proccesing*. Fourth Edition. Malaysia: Global Edition. Peaeson Education Limited, 2019.
- [11] *TuDashboard*. URL: <https://tudashboard.com/grafico-de-mapa-de-calor/>.



- [12] Roberto Depaoli y Luis A. Fernández. «Optimización de la ecualización del histograma en el procesamiento de imágenes digitales». Departamento de Ingeniería e Investigaciones Tecnológicas. Universidad Nacional de La Matanza, mar. de 2012.
- [13] David Herrera Sánchez. «Construcción automática de características para segmentación de imágenes de pulmones afectados por COVID-19 utilizando programación genética». Instituto de Investigaciones en Inteligencia Artificial. Universidad Veracruzana, mayo de 2021.
- [14] Sergio Sada Pezonaga. «Tratamiento de imágenes con ruido impulsivo mediante reglas difusas y algoritmos genéticos». E.T.S. de Ingeniería Industrial, Informática y Telecomunicaciones, 2015.



A. Tabla 11

#	Tomografía	$\sigma_{\text{ruidooscuro}}$	$\mu_{\text{ruidooscuro}}$	χ^2
1	2020.01.24.919183 _{p27} 132.png	0.010086	0.166177	20.536829
2	2020.02.23.20026856 – p17 – 115 ₃ .png	0.011654	0.221842	2.298203
3	bmj.m606.full – p4 – 22 ₇ .png	0.005220	0.107493	4.518013
4	2020.01.24.919183 – p27 – 133.png	0.009701	0.165853	91.064755
5	2020.02.10.20021584 – p6 – 52 ₃ .png	0.014072	0.207994	4.015357
6	2020.03.18.20038125 – p16 – 56 – 2.png	0.002612	0.222364	29.105965
7	chapter – 17 – p10 – 515.png	0.013391	0.132372	24.862708
8	2020.02.10.20021584 – p6 – 52 ₁ .png	0.010059	0.206409	0.783227
9	2020.02.11.20022053 – p12 – 67 ₂ .png	0.014679	0.089513	31.475878
10	2020.02.11.20022053 – p12 – 67 ₃ .png	0.001034	0.070407	9.444444
11	2020.02.23.20026856 – p17 – 115 ₁ .png	0.015426	0.222261	3.449891
12	2020.02.23.20026856 – p17 – 115 ₅ .png	0.022126	0.213323	2.283519
13	bmj.m606.full – p4 – 22 ₂ .png	0.005459	0.109984	14.930466
14	bmj.m606.full – p4 – 22 ₆ .png	0.004143	0.110180	8.601337
15	bmj.m606.full – p4 – 22 ₃ .png	0.006153	0.109644	14.886075
16	2020.02.24.91983 – p27 – 134.png	0.008930	0.164982	104.890730
17	2020.02.24.91983 – p27 – 135.png	0.013043	0.168026	147.795561
18	2020.02.10.20021584 – p6 – 52 ₇ .png	0.006229	0.214013	4.258686
19	2020.02.10.20021584 – p6 – 52 ₁ 1.png	0.015698	0.226988	7.425717
20	2020.02.10.20021584 – p6 – 52 ₁ 5.png	0.031356	0.229667	5.652414
21	2020.02.10.20021584 – p6 – 52 ₄ .png	0.005540	0.213704	12.271349
22	2020.02.10.20021584 – p6 – 52 ₈ .png	0.007438	0.208513	7.074792
23	2020.03.22.20034041 – p17 – 80 – 2.png	0.009797	0.169242	83.219511
24	2020.02.10.20021584 – p6 – 52 ₁ 5.png	0.002820	0.211681	2.142830
25	2020.02.11.20021493 _{p16} 109 ₂ .png	0.013304	0.207085	16.611718
26	2020.02.23.20026856 – p17 – 115 ₄ .png	0.011543	0.225285	2.387724
27	2020.02.24.20027052 – p8 – 73 ₁ .png	0.061284	0.016221	1.887654
28	2020.03.03.20030775 – p10 – 88 ₁ .png	0.082276	0.182761	9.936236
29	2020.03.03.20030775 – p12 – 93 ₀ .png	0.043971	0.167063	28.688839
30	2020.03.21.20040691 – p18 – 67 – 3.png	0.008132	0.227589	12.177675
	Promedio	0.015584	0.172949	23.622605



B. Tabla 12

#	Tomografía	$\sigma_{\text{ruidobrillo}}$	$\mu_{\text{ruidobrillo}}$	χ^2
1	2020.01.24.919183 _{p27} 132.png	0.004525	0.994263	265.946977
2	2020.02.23.20026856 – p17 – 115 ₃ .png	0.018399	0.958069	1.694923
3	bmj.m606.full – p4 – 22 ₇ .png	0.086779	0.799465	10.307200
4	2020.01.24.919183 – p27 – 133.png	0.003841	0.991438	226.262082
5	2020.02.10.20021584 – p6 – 52 ₃ .png	0.014402	0.872347	7.298012
6	2020.03.18.20038125 – p16 – 56 – 2.png	0.043103	0.882955	134.094454
7	chapter – 17 – p10 – 515.png	0.073741	0.712714	1.091899
8	2020.02.10.20021584 – p6 – 52 ₁ .png	0.040897	0.824680	52.038577
9	2020.02.11.20022053 – p12 – 67 ₂ .png	0.022618	0.899028	50.589820
10	2020.02.11.20022053 – p12 – 67 ₃ .png	0.040009	0.879636	91.260545
11	2020.02.23.20026856 – p17 – 115 ₁ .png	0.017504	0.944358	7.203709
12	2020.02.23.20026856 – p17 – 115 ₅ .png	0.016074	0.944044	8.180834
13	bmj.m606.full – p4 – 22 ₂ .png	0.051769	0.824812	68.796132
14	bmj.m606.full – p4 – 22 ₆ .png	0.043481	0.839076	16.126335
15	bmj.m606.full – p4 – 22 ₃ .png	0.068764	0.836175	40.280911
16	2020.02.24.91983 – p27 – 134.png	0.005938	0.994336	516.492236
17	2020.02.24.91983 – p27 – 135.png	0.030840	0.931863	163.385987
18	2020.02.10.20021584 – p6 – 52 ₇ .png	0.013061	0.874301	54.391500
19	2020.02.10.20021584 – p6 – 52 ₁₁ .png	0.006769	0.877338	42.286579
20	2020.02.10.20021584 – p6 – 52 ₁₅ .png	0.017541	0.857900	94.053898
21	2020.02.10.20021584 – p6 – 52 ₄ .png	0.012266	0.787288	23.091284
22	2020.02.10.20021584 – p6 – 52 ₈ .png	0.041502	0.843416	52.920750
23	2020.03.22.20034041 – p17 – 80 – 2.png	0.013602	0.911824	78.471417
24	2020.02.10.20021584 – p6 – 52 ₁₅ .png	0.039786	0.896058	32.509662
25	2020.02.11.20021493 _{p16} 109 ₂ .png	0.017273	0.862376	28.123907
26	2020.02.23.20026856 – p17 – 115 ₄ .png	0.019032	0.849278	9.466359
27	2020.02.24.20027052 – p8 – 73 ₁ .png	0.022835	0.397257	13.490955
28	2020.03.03.20030775 – p10 – 88 ₁ .png	0.037254	0.796020	16.215371
29	2020.03.03.20030775 – p12 – 93 ₀ .png	0.030989	0.565203	22.614485
30	2020.03.21.20040691 – p18 – 67 – 3.png	0.056996	0.902705	45.470555
	Promedio	0.030386	0.851674	71.805245



C. Código: ImageToHisto()

```

#include "TCanvas.h"
#include "TH1.h"
#include "TAxis.h"
#include "TStyle.h"
#include "TSystem.h"
#include "TColor.h"
#include <string.h>
#include <iostream>
#include <ostream>
#include <cstdio>
#include <fstream>
#include "TF2.h"
#include "TExec.h"
using namespace std;
void ImageToHisto()
{
    gStyle->SetOptTitle(1);           /**Para visualizar el titulo 0; 1
    gStyle->SetOptStat(1);           /**Vizualizar leyenda histograma 0;
        1
    gStyle->SetCanvasColor(10);       /**Color del fondo grafica
    gStyle->SetCanvasBorderMode(0);   /**Margen de toda la grafica
    gStyle->SetFrameLineWidth(1);     /**Ancho del marco de la grafica
    gStyle->SetFrameFillColor(kWhite); /**Color del fondo interno
    gStyle->SetPadTickX(1);           /** Vizualizar numeros en el x
        superior
    gStyle->SetPadTickY(1);           /** Vizualizar numeros en el y
        superior
    gStyle->SetPadBottomMargin(0.07); /** Para ajustar margen inferior
    gStyle->SetPadLeftMargin(0.08);   /** Ajustar la imagen a la
        izquierda
    gStyle->SetPadTopMargin(0.08);    /** Ajustar margen superior
    gStyle->SetPadRightMargin(0.09);  /** Ajustar la imagen a la
        izquierda
    gStyle->SetLineWidth(2);          /**Ancho del cuadrado de la leyenda
    gStyle->SetLabelSize(0.04, "xyz"); /**Numeros de los ejes

```



```

gStyle->SetLabelOffset(0.01,"y"); /**Aleja los numeros del eje y
gStyle->SetLabelOffset(0.01,"x"); /**Aleja los numeros del eje x
gStyle->SetLabelColor(kBlack,"xyz"); /**Color de los numeros de los
    ejes
gStyle->SetTitleOffset(1.1,"y"); /**Alejar leyenda del eje y
gStyle->SetTitleOffset(1.2,"x"); /**Alejar leyenda del eje x
gStyle->SetTitleFillColor(kWhite); /**Color del fondo del titulo
gStyle->SetLegendBorderSize(1); /**Margen de la leyenda**//
gStyle->SetLegendFillColor(kWhite); /**Color de relleno de la leyenda
    **//
gStyle->SetLegendFont(62);

////////////////////////////////////
///// Initial parameters /////
////////////////////////////////////

const Int_t boxes=3;
std::ofstream outFile("2020.01.24.919183_p27_132.dat",std::ios::app);
/**(ubicacion donde se desea guardar el archivo .dat)**//
TASImage image("2020.01.24.919183_p27_132.png");
/**(ubicacion de la imagen"/nombre de la imagen)**//

////////////////////////////////////
///// Image digitalization /////
////////////////////////////////////

/**obtencion del ancho y alto de la imagen
const UInt_t yPixels = image.GetHeight();
const UInt_t xPixels = image.GetWidth();
UInt_t *argb = image.GetArgbArray();

/** Eje X y eje Y
Double_t X[xPixels];
Double_t Y[yPixels];
const Int_t n = yPixels*xPixels;
Float_t Grey[xPixels][yPixels];
cout<<xPixels<<" "<<yPixels<<endl;

```



```

X[0]=-1; // Inicializar el conteo de los ejes
Y[0]=1;
Double_t XX,YY;

// Este es el incremento que va teniendo de un ponto del eje X o Y a otro
Double_t incx=2.0/xPixels;
Double_t incy=2.0/yPixels;

// Se define el histograma de la imagen exportada
TH2D* h = new TH2D( "h", "RedigitalizacioneunaTACdetorax , MasculinoCOVID
-19", xPixels , -1 , 1, yPixels , -1, 1 );

Double_t maxA=0.;
Double_t minA=50.;
////////////////////////////////////
////////// Construccion de la imagen exportada //////////
////////////////////////////////////
for(int j=1;j<xPixels; j++) //Llenado cada componente del eje X y eje Y
{
    X[j]=X[0]+j*incx;
}
for(int j=1;j<yPixels; j++)
{
    Y[j]=Y[0]-j*incy;
}
Double_t cordx;
Double_t cordy;
Double_t x,y;
for (int row=0; row<xPixels; row++) // Ciclo filas la matriz
    intensidades
{
    for (int col=0; col<yPixels; col++) //Ciclo columnas matriz
        intensidades
    {
        int index = col*xPixels+row;
        float grey = float(argb[index]&0xff)/256;
    }
}

```



```

    h->SetBinContent(row, yPixels-col, pow(grey,1)); //se llena el
        histograma (fila, columna, intensidad elevados a potencia 1)
    Grey[row][col]=grey; // Se llena la matriz de intensidades
    outFile <<X[row]<<" "<<Y[col]<<" "<<Grey[row][col]<<endl;
}
}
cout<<"TotalSegment:"<<cordx<<" "<<cordy<<" "<<x<<" "<<y<<" "<<maxA<<endl;
TCanvas *c1 = new TCanvas("c1", "c1", 0, 0, 600, 400);
gStyle->SetPalette(53);
Double_t binmax = h->GetMaximum();
//Double_t binmin = h->GetMinimum();
h->SetMaximum(1);          /**Establece el valor maximo de la paleta
h->SetContour(99);        /**Nmero de divisiones de la paleta
h->SetStats(1);
h->Draw("COLZ");
return 0;
}

```



D. Código: gauss medias()

```

void gauss_medias ()
{ /*
    cout<< " ... Setting Style of plots ..." <<endl;
    gStyle->Reset(" Plain ");
    gStyle->SetOptTitle(0);
    gStyle->SetOptStat(0000);
    gStyle->SetCanvasColor(10);
    gStyle->SetCanvasBorderMode(0);
    gStyle->SetFrameLineWidth(1);
    gStyle->SetFrameFillColor(kWhite);
    gStyle->SetPadColor(10);
    gStyle->SetPadTickX(1);
    gStyle->SetPadTickY(1);
    gStyle->SetPadBottomMargin(0.12);
    gStyle->SetPadLeftMargin(0.1);
    gStyle->SetPadTopMargin(0.03);
    gStyle->SetPadRightMargin(0.03);
    gStyle->SetHistLineWidth(0.5); // gStyle->SetHistLineWidth(1);
    gStyle->SetHistLineColor(kRed);
    gStyle->SetFuncWidth(2);
    gStyle->SetFuncColor(kGreen);
    gStyle->SetLineWidth(2);
    gStyle->SetLabelSize(0.04," xyz ");
    gStyle->SetLabelOffset(0.01," y");
    gStyle->SetLabelOffset(0.01," x");
    gStyle->SetLabelColor(kBlack," xyz ");
    gStyle->SetTitleSize(0.05," xyz "); //0.05 Lebel on Y
    gStyle->SetTitleOffset(0.85," y"); //0.85
    gStyle->SetTitleOffset(1.2," x"); //0.95
    gStyle->SetTitleFillColor(kWhite);
    gStyle->SetTextSizePixels(28); //26
    gStyle->SetFont(42); // 42

    gStyle->SetLegendFillColor(kWhite);
    gStyle->SetLegendFont(62); */

```



```

gStyle->SetLegendBorderSize(0); //Para que la leyenda de la campana de
    Gauss vaya sin contorno
//TH2D* h = new TH2D( "h", "No Covid digitalization", 1024, -1 , 1, 1024,
    -1, 1); //lo vamos a trabajar en una sola dimension (1D)
//TH1* h = new TH1F("h", "Paciente Masculino con opacidades circulares
    COVID-19", 50, 0.6, 1.0);
TH1F* h = new TH1F("h", "Ruido oscuro", 18, 0.0, 0.40);
ifstream File("mediaoscura.dat"); // .dat Nombre de archivo
//TCanvas *c1 = new TCanvas("", "Canvas", 600, 600);

{
    Double_t w = 600;
    Double_t height = 600;
    TCanvas * c1 = new TCanvas("c", "c", w, height);
    c->SetWindowSize(w + (w - c->GetWw()), height + (height - c->GetWh()));
}

Double_t X1, X2, X3;

while (File >> X1 && File >> X2 && File >> X3)
{
    // if(0.0 < X1 && X1 < 0.5)
        // {
            // if(-sqrt(0.0225 - X1*X1) < X2 && X2 < sqrt(0.0225 - X1*X1))
                // {
                    h->Fill(X3);
                // }
            // }
}

h->GetXaxis()->SetTitle("Intensidad del bin");
h->GetXaxis()->SetTitleSize(0.04);
h->GetXaxis()->CenterTitle(true);

h->GetYaxis()->SetTitle("N#acute{u}merodeentradas");
h->GetYaxis()->SetTitleSize(0.04);
h->GetYaxis()->SetTitleOffset(0.8); //para alejar la leyenda del eje
h->GetYaxis()->CenterTitle(true);

```



```

h->Draw ();

Double_t par[3];
TF1 *gauss= new TF1("gauss", "gaus", 0.05, .35);
gauss->SetLineColor(3);
gauss->SetLineWidth(6); //Grosor de la campana de Gauss
gauss->SetParName(0,"constant");
gauss->SetParName(1,"mean");
gauss->SetParName(2,"sigma");

// gStyle->SetOptStat(0);
//gStyle->SetOptFit(111);
h->Fit(gauss,"R");
h->SetLineColor(2);
h->SetLineWidth(6); //Grosor de los histogramas
h->Draw("");

gauss->GetParameters(par);
gauss->SetParameters(par);
Float_t sigma = par[2];
Float_t mean = par[1];
Double_t Chi2 = gauss->GetChisquare();
Int_t Entries = h->GetEntries();
Double_t Errorsigma = gauss->GetParError(2);
Double_t Errormean = gauss->GetParError(1);
Double_t NDF = gauss->GetNDF();
Double_t Chi2NDF = Chi2/NDF;

cout<<Chi2<<" , "<<NDF<<" , Chisquare / ndf : "<<Chi2NDF<<endl;

TLegend *leg=new TLegend(0.5,0.65,0.80,0.85,"");
leg->SetTextFont(72);
leg->SetTextSize(0.030);

leg->AddEntry(h,"Datos","l");

```



```
leg->AddEntry(gauss,"Funci#acute{o}ndeGauss","1");
leg->AddEntry((TObject*)0, Form("Entries=%i",Entries), "");
leg->AddEntry((TObject*)0, Form("#sigma=%f#pm%f",sigma,Errorsigma), "");
leg->AddEntry((TObject*)0, Form("mean=%f#pm%f",mean,Errormean), "");
leg->AddEntry((TObject*)0, Form("#Chi^{2}/ndf=%f",Chi2NDF), "");
leg->Draw();
}
```