

BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA



BUAP

TESIS DE MAESTRÍA

Detección de COVID-19 por Medio de XR y CT de Tórax con Aprendizaje Profundo

Presenta:

Víctor Daniel Machorro García

Director:

Dr. Ivo Humberto Pineda
Torres

Asesores:

Dra. María Josefa Somodevilla
García
Dr. Manuel Isidro Martín Ortiz

Tesis para obtener el grado de Maestro en Ciencias de la Computación en la

Facultad de Ciencias de la Computación

21 de enero de 2022

BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

Resumen

Facultad de Ciencias de la Computación

Maestro en Ciencias de la Computación

Detección de COVID-19 por Medio de XR y CT de Tórax con Aprendizaje Profundo

por Víctor Daniel Machorro García

El SARS-CoV-2 el coronavirus recientemente descubierto en diciembre de 2019 en la región de Wuhan China ha provocado un gran impacto en la salud, economía, sociedad y en general la mayoría de los aspectos que la vida diaria de todo el mundo. El mas importante de ellos es la salud y vida de las personas, al momento de la realización de este trabajo no existe aun un tratamiento científicamente comprobado y validado por la organización mundial de la salud (WHO por sus siglas en ingles). El virus provoca la enfermedad denominada COVID-19 (Coronavirus Disease 2019) la cual tiene manifestaciones, en la mayoría de los infectados, como fiebre, diarrea, dolor de garganta, tos, dolor de cuerpo, dolor de cabeza y en los casos mas graves dificultad para respirar causada por una neumonía asociada al virus. Esta es la principal causa de muerte por COVID-19, las personas que resultan mayormente afectadas pertenecen al grupo de mayores de 60 años, mujeres embarazadas y/o con alguna comorbilidad como obesidad, diabetes, hipertensión, EPOC o sistema inmune deprimido. En los casos mas extremos es necesario realizar un procedimiento invasivo en el paciente donde se le coloca un respirador artificial. Debido a la ausencia de un tratamiento, así como de una vacuna masivamente disponible los sistemas de salud de la mayoría de los países se han visto superados en su capacidad de atención provocando así una cantidad mayor de fallecimientos.

La principal forma de detección del virus es por medio de una prueba PCR, la cual puede tardar entre 24 y 48 horas, en los sistemas de salud saturados estos resultados pueden no estar disponibles y puede que su resultado tarde aún mas. Un mecanismo que se utiliza para obtener un resultado mas rápido en pacientes que presentan alguna afectación en los pulmones es la detección de la enfermedad por medio de imágenes de rayos X y tomografía computarizada de tórax. Las imágenes de estos cuentan con características identificables por personal medico calificado.

Este proyecto utilizo la inteligencia artificial y el aprendizaje profundo para la clasificación automática de estas imágenes, ademas se genero un conjunto de datos propio que pudiera tener una representación de la población Mexicana, uno de los principales problemas al inicio del proyecto fue la ausencia de datos públicos validos y fiables para el entrenamiento de la red neuronal. Después de la generación de un conjunto de datos diverso, robusto y relativamente grande para el proyecto, se implemento y entreno una red neuronal convolucional de aprendizaje profundo bajo la plataforma de desarrollo de Tensorflow en el lenguaje de programación Python. Se realizaron diversas pruebas comparativas con otros modelos utilizados generalmente en imágenes médicas como lo son VGG, Inception, Xception y RESNET 50, teniendo como resultado un sistema eficiente, confiable y robusto, alcanzando niveles de precisión por encima del 90 %, aun así este sistema como muchos otros estudiados, debe ser utilizado siempre bajo la supervisión y validación de personal correctamente capacitado el cual deberá otorgar el diagnostico final, las características y resultados del sistema permitiría aliviar la carga de trabajo que tiene el personal de salud cuando los sistemas de salud se saturan.

El uso de aprendizaje profundo para el tratamiento de imágenes medicas es un área que continua en desarrollo en gran parte a la nueva tecnología respecto a los aceleradores gráficos de propósito general GPGPU los cuales permiten obtener resultados en fracciones de tiempo respecto a los procesadores tradicionales, gracias al uso del equipo IBM AC922 con tarjetas NVidia V100 se lograron realizar múltiples experimentos que no hubiera sido posible realizar en equipos de computo convencionales.

Gracias al la combinación de las herramientas de aprendizaje profundo, el equipo

de cómputo de alto desempeño y de un conjunto de datos robusto se lograron obtener resultados que logran detectar en un porcentaje satisfactorio al conjunto de personas con afectaciones en los pulmones causadas por COVID-19 y sobre todo logra descartar a personas que no tienen afectaciones por esta enfermedad logrando así el principal objetivo que es ayudar al personal médico en la labor de detección y clasificación de pacientes.

Agradecimientos

En primer lugar quisiera agradecer a mis tutores los Doctores Ivo Humberto Pinéda Tórres, Manuel Isidro Martín Ortiz y la Doctora Maria Josefa Somodevilla, por todo su apoyo y dedicación hacia mi persona así como a este trabajo.

También agradecer a la Benemérita Universidad Autónoma de Puebla y al Postgrado de la Facultad de Ciencias de la Computación por hacer posible la realización de esta investigación.

Los autores agradecen al Laboratorio Nacional de Supercómputo del Sureste de México (LNS), perteneciente al padrón de laboratorios nacionales CONACYT, por los recursos computacionales, el apoyo y la asistencia técnica brindados, a través del proyecto No.202101032C .

Índice general

Resumen	IV
Agradecimientos	VII
1. Introducción	1
1.1. COVID-19	1
1.2. Las Imágenes Rx y CT	1
1.2.1. Radiografías (Rx)	2
1.2.2. Tomografías Computarizada(CT)	2
1.3. Aprendizaje Profundo en Imágenes Médicas	3
1.3.1. Introducción	3
1.3.2. Aprendizaje supervisado y no supervisado	4
1.3.3. Redes Neuronales Profundas Convolucionales	5
1.3.4. Aplicaciones en Radiología	6
Clasificación de Imágenes	6
Detección de Objetos	6
Segmentación de Imágenes	7
Generación de Imágenes	7
Transformación de Imágenes	8
2. Capítulo 2- Marco Conceptual y Contextual	9
2.1. Introducción	9
2.2. Manifestación Clínica	10
2.3. Pruebas Diagnosticas	11
2.4. Tratamiento de la Enfermedad	12
2.5. Perfil de Una afectación en los Pulmones Causada por COVID-19	13
Perfil de CT de Torax en Pacientes con COVID-19	14
Perfil de XR de Torax en Pacientes con COVID-19	16
3. Capítulo 3 - Marco Metodológico	19
3.1. Conjunto de Datos	19
3.1.1. introducción	19
3.1.2. Datos Públicos de Imágenes de Tórax de Pacientes con COVID-19	20
CheXpert X-Ray Dataset	20
covid-chestxray-dataset COVID-19 image data collection	21
COVID-CT-Dataset	22
Padchest-xray-dataset	22
Actualmed COVID-19 Chest X-ray Dataset Initiative	22
COVID-CTset	22

	COVID-19 RADIOGRAPHY DATABASE	22
3.1.3.	Creación de Conjunto de Datos Propio	23
	Características de las Imágenes	23
	Procesamiento de imágenes en formato DICOM	23
3.1.4.	Conjunto de Datos Final	25
3.1.5.	Preprocesamiento de las Imágenes Médicas	29
3.2.	Modelo de Red Neuronal Artificial Profunda	34
3.2.1.	Introducción	34
3.2.2.	Implementación de modelos de red existentes	34
	VGG	34
	RESNET	35
	InceptionV3	35
	Xception	36
3.2.3.	Diseño	37
3.2.4.	Implementación	37
	Hardware	37
	Software	40
3.2.5.	Entrenamiento	40
4.	Capítulo 4 - Análisis de Resultados	43
4.1.	Introducción	43
4.2.	Métricas de Evaluación	43
4.2.1.	Exactitud	44
4.2.2.	Precisión	44
4.2.3.	Sensibilidad y la Especificidad	45
4.2.4.	ROC	45
4.3.	Resultados del modelo implementado	46
4.3.1.	Matriz de Confusión	46
4.3.2.	Reporte de Clasificación	46
4.3.3.	la Curva ROC	48
4.4.	Comparativa con los modelos Inception, VGG, ResNet, Xception	48
5.	Capítulo 5 - Conclusiones	53
5.1.	Conclusiones	53
	Bibliografía	57

Índice de figuras

1.1. ANN y DNN	4
1.2. Red Neuronal Profunda Convolutacional	5
2.1. Imagen CT de Tórax Coronal. Se muestra las opacidades de tipo vidrio deslustrado con morfología redonda.	14
2.2.	15
2.3.	16
2.4. Paciente con daño avanzado en los pulmones y consolidación bilateral. . .	17
3.1. Distribución de las imágenes de pacientes con neumonía del conjunto de datos covid-chestxray-dataset.	21
3.2. Adquisición de imágenes de archivos DICOM	24
3.3. Histograma de imagen sin HU	25
3.4. Histograma de imagen con HU	26
3.5. Planos Axiales Sin Normalización	27
3.6. Planos Axiales Con Normalización	28
3.7. Radiografía de Tórax original sin filtro CLAHE	30
3.8. Radiografía de Tórax con filtro CLAHE aplicado	31
3.9. Preprocesamiento de Región de Interés	32
3.10. Generación sintética de imágenes	33
3.11. Modelo completo de procesamiento y clasificación	38
3.12. Modelo de red CNN propuesto	39
4.1. Matriz de Confusión para Nuestro modelo con CT	47
4.2. ROC para el modelo entrenado con Tomografías	50

Índice de cuadros

3.1. Unidades Hounsfield	24
3.2. Conjunto de Datos para entrenamiento de la Red.	29
4.1. Referencia de traducción de términos relacionados con las métricas utilizadas en este proyecto.	44
4.2. Reporte de Clasificación para el modelo propio con datos CT	48
4.3. Reportes de Clasificación de los modelos comparativos.	49
4.4. Matrices de Confusión de los modelos de Red comúnmente utilizados en imágenes, utilizando tomografías	51
4.5. Curvas ROC de modelos de Red comúnmente utilizados en imágenes, utilizando tomografías	52

Lista de Abreviaturas

CT	Computed Tomography
CXR	Chest X-ray
MERS	Middle East Respiratory Syndrome
SARS	Severe Acute Respiratory Syndrome
COVID-19	COrona VIRus Disease 2019
NIfTI	Neuroimaging Informatics Technology Initiative
DICOM	Digital Imaging and Communications in Medicine
JPG, JPEG	Joint Photographic Experts Group
PNG	Portable Network Graphics
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
GAN	Generative Adversarial Network
GPGPU	General Porpouse Graphics Processor Unit
HU	Hounsfield Unit
CLAHE	Contrast Limited Adaptative Histogram Equalization
GAN	Generative Adversarial Network

Capítulo 1

Introducción

1.1. COVID-19

La enfermedad COVID-19 (Coronavirus Disease 2019) causada por el virus SARS-CoV-2 se presentó a finales del año 2019 en China, para inicios del año 2020 ya se había esparcido a gran parte de Asia y Europa poco después llegó a América. Esta enfermedad afecta principalmente a los grupos de personas que tienen alguna comorbilidad o sistema inmune deprimido provocando en ellos una neumonía, aun no se cuenta con un tratamiento efectivo una vez que los pacientes han sido infectados, en el curso de este proyecto se han desarrollado varias vacunas que permiten obtener cierto nivel de protección especialmente antes los síntomas graves, la mayoría de ellas sobrepasa el 95 por ciento de efectividad en este caso y todas ellas previene la sintomatología en más del 50 por ciento de los vacunados. Con la vacuna se vislumbra un panorama mejor para las condiciones de vida en general de las personas permitiendo que gradualmente y en un tiempo relativamente corto se pudiera retomar las actividades no esenciales que se vieron limitadas por la pandemia, aun así existe el problema de la aplicación universal debido a la escasez en los materiales, fábricas, medios de distribución y mano de obra para la elaboración de las vacunas. México se encuentra en una situación complicada debido a la distribución de su población, la escasez de vacunas y la cantidad de personas con comorbilidades, es el país con el índice de obesidad infantil más grande y en personas adultas se encuentra en segundo lugar. El COVID-19 es muy probable que permanezca como una enfermedad endémica por lo que es importante contar con un sistema de vacunación periódico y medios que permitan contar con una detección temprana de la enfermedad para prevenir afectaciones o incluso la muerte de los enfermos.

1.2. Las Imágenes Rx y CT

Las técnicas de diagnóstico por imágenes ayudan a reducir las causas de una lesión o una enfermedad, y garantizan que el diagnóstico sea preciso. Estas técnicas incluyen las radiografías, las tomografías computarizadas (TC) y las imágenes por resonancia magnética (IRM).

Estas herramientas de diagnóstico por imágenes le permiten al médico "ver" el interior del cuerpo a fin de obtener una imagen de los huesos, los órganos, los discos, los meniscos, los músculos, los tendones, los nervios y los cartílagos. De esta forma, el médico puede determinar si existe alguna anomalía.

Existe una gran cantidad de enfermedades o lesiones que pueden ser diagnósticas por medio de este tipo de imágenes, algunas de ellas son cáncer, fracturas, trombos,

afectaciones cardíacas y sobre las que trabaja este proyecto las lesiones en los pulmones, en específico la neumonía asociada a COVID-19. Es un tema que presenta grandes retos debido a que en específico esta manifestación de la enfermedad podría ser confundida con otra que cause neumonía, entre las principales neumonías que se tiene que diferenciar son las causadas por los virus MERS, SARS-COV-1, Influenza y las causadas por diversos tipos de bacterias.

1.2.1. Radiografías (Rx)

Las radiografías (rayos X) son la técnica de diagnóstico por imágenes más común y más accesible. Los rayos X son una forma de radiación electromagnética, similares a la luz visible. Sin embargo, a diferencia de la luz, los rayos X tienen una mayor energía y pueden pasar a través de la mayoría de los objetos, incluyendo el cuerpo. Los rayos X médicos se utilizan para generar imágenes de los tejidos y las estructuras dentro del cuerpo. Si los rayos X que viajan a través del cuerpo también pasan a través de un detector de rayos X al otro lado del paciente, se formará una imagen que representa las sombras formadas por los objetos dentro del cuerpo.

Un tipo de detector de rayos X es la película fotográfica, aunque existen muchos otros tipos de detectores que se utilizan para producir imágenes digitales. Las imágenes de rayos X que resultan de este proceso se llaman radiografías.

La parte del cuerpo a la que le toman las imágenes se ubica entre la máquina de rayos X y la película fotográfica o sensor de rayos X digital. Debe mantenerse inmóvil mientras la máquina envía brevemente ondas electromagnéticas (radiación) a través del cuerpo, exponiendo la película para que refleje su estructura interna. El nivel de exposición a la radiación de los rayos X no es perjudicial.

Los huesos, las calcificaciones, algunos tumores y otra materia densa aparecen en color blanco o claro debido a que absorben la radiación. Los tejidos blandos menos densos y las fracturas en los huesos permiten que la radiación los traspase, lo que hace que estas partes se vean más oscuras en la película de rayos X.

En algunas circunstancias, se puede inyectar un material de contraste o tinte en una articulación mientras se toman las radiografías. Este procedimiento, que se llama *artrografía*, ayuda a delinear las estructuras de tejido blando en la articulación. También puede confirmar la ubicación de la aguja en la articulación cuando se extrae el líquido o se inyecta medicamento en la articulación.

Es posible que las radiografías no muestren muchos detalles como una imagen producida con técnicas más sofisticadas. Sin embargo, son la herramienta de diagnóstico por imágenes más común que se utiliza para evaluar un problema ortopédico y están fácilmente disponibles en la mayoría de los centros médicos.

1.2.2. Tomografías Computarizada(CT)

La tomografía computarizada (CT) es una herramienta de diagnóstico por imágenes que combina rayos X con tecnología computarizada para producir una imagen más detallada y transversal del cuerpo. Una CT le permite al médico ver el tamaño, la forma y la posición de las estructuras que están en la profundidad del cuerpo, como órganos, tejidos o tumores. El término tomografía computarizada, o CT, se refiere a un procedimiento computarizado de imágenes por rayos X en el que se proyecta un haz angosto de rayos X a un paciente y se gira rápidamente alrededor del cuerpo, produciendo

señales que son procesadas por la computadora de la máquina para generar imágenes transversales o cortes del cuerpo. Estos cortes se llaman imágenes tomográficas y contienen información más detallada que los rayos X convencionales. Una vez que la computadora de la máquina recolecta varios cortes sucesivos, se pueden apilar digitalmente para formar una imagen tridimensional del paciente que permita más fácilmente la identificación y ubicación de las estructuras básicas, así como de posibles tumores o anomalías.

Para este examen, el paciente debe recostarse y permanecer inmóvil tanto como sea posible sobre una camilla que se desliza hacia el centro de un escáner de CT, similar a un cilindro. El proceso es sin dolor. Un tubo de rayos X rota lentamente alrededor del cuerpo, tomando varias imágenes desde todas las direcciones. Una computadora combina las imágenes para producir una vista clara y bidimensional en una pantalla de televisión.

Una CT es más costosa y demora más tiempo que una radiografía común. Se puede realizar en un hospital o en un centro de diagnóstico por imágenes en forma ambulatoria.

1.3. Aprendizaje Profundo en Imágenes Médicas

1.3.1. Introducción

La tecnología de la inteligencia artificial, con la ayuda del poder de cómputo avanzado, una gran cantidad de datos y nuevos algoritmos se convierte en una de las herramientas más populares y poderosas que existen en la actualidad. Se ha aplicado a diversos campos de estudio como por ejemplo la salud, industria, vida cotidiana solo por mencionar algunos. La inteligencia artificial se puede dividir en 3 grandes categorías. Una de ellas es el enfoque simbólico el cual entrega resultados utilizando una máquina de búsqueda basada en reglas. Otra es el enfoque basado en los teoremas Bayesianos. La última es en enfoque basado en el conexionismo de redes neuronales profundas. Mientras que cada una tiene sus ventajas y desventajas el enfoque de redes neuronales profundas está ganando atención para resolver problemas de alta complejidad.

El aprendizaje máquina es un subconjunto de la inteligencia artificial que aprende de los datos por sí mismo con una mínima intervención humana para realizar una clasificación o predecir condiciones a futuro o inciertas. Murphy, 2013 algunas de las tareas comunes del aprendizaje máquina son la regresión, clasificación, detección y segmentación por mencionar algunas. Por lo general los conjuntos de datos del aprendizaje máquina consisten en conjuntos de entrenamiento, validación y prueba. El conjunto de entrenamiento sirve para aprender las características que son de interés y el conjunto de validación permite validar este aprendizaje. Finalmente uno puede confirmar la precisión por medio del conjunto de prueba.

Una red neuronal artificial (ANN por sus siglas en inglés) es un algoritmo inspirado en el funcionamiento del cerebro humano donde existen un conjunto de capas con nodos conectados, esta compuesta de capas de entrada, intermedias u ocultas y de salida. Durante el entrenamiento cada nodo es parametrizado con pesos los cuales van cambiando, dependiendo de la implementación que se tenga este proceso se detiene al alcanzar cierta precisión o al llegar a una cantidad de iteraciones. Sin embargo las ANN tienen limitaciones sobre todo en un sobre ajuste al conjunto de entrenamiento lo

que significa que solo funcionara correctamente para ese conjunto de datos perdiendo generalidad. El un progreso relativamente reciente se han expandido las ANN en redes neuronales artificiales de aprendizaje profundo (DNN por sus siglas en ingles) las cuales permiten incrementar de manera substancial las capas internas que conectan las capas de entrada y de salida. Las DNN generalmente muestran mejores resultados en las áreas de clasificación y regresión (He y col., 2016). Cada capa de la DNN optimiza su resultado en una maquina sin supervisión y restringida de Boltzmann para prevenir un mínimo local o problemas de sobre entrenamiento. 1.1

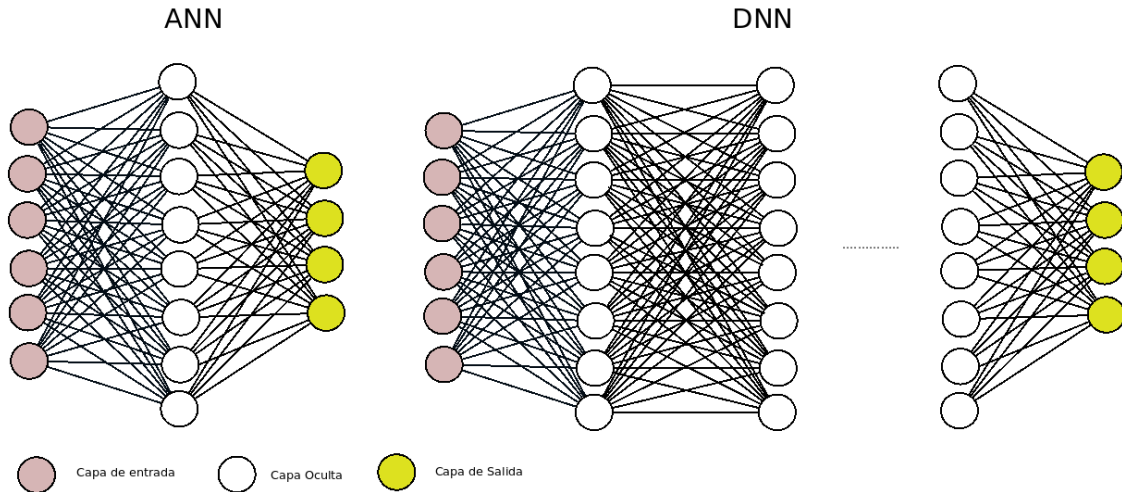


FIGURA 1.1: ANN y DNN

Recientemente se han utilizado con mayor frecuencia los algoritmos de aprendizaje profundo para resolver diversos problemas en imágenes médicas. Gracias a esta tecnología estos problemas han obtenido un resultado exitoso permitiendo que el personal médico dedique su tiempo a otras tareas. Sin embargo esta tecnología no significa el remplazo del personal médico, en especial a los imagenólogos, en cambio se utiliza como un auxiliar en el diagnostico.

1.3.2. Aprendizaje supervisado y no supervisado

Los algoritmos de aprendizaje maquina pueden ser categorizados en supervisado, no supervisado y aprendizaje reforzado. El aprendizaje reforzado no es adecuado en aplicaciones medicas debido a que la decisión del algoritmo afectaría la salud futura del paciente así como a la opción del tratamiento, como resultado los efectos a largo plazo son difíciles de estimar (Gottesman y col., 2019). La principal diferencia entre los sistemas supervisados y no supervisados radica en que en el primero los datos en el conjunto de entrenamiento tiene una etiqueta que los identifica la clase a la que pertenecen. El aprendizaje supervisado infiere una relación matemática entre las entradas y las salidas etiquetadas mientras que el no supervisado infiere una función que expresa características ocultas que residen en los datos de entrada.

En el aprendizaje supervisado los datos de salida pueden tener un valor categórico o un valores numérico continuo, se convierte en un problema de clasificación o de reconocimiento de patrones cuando el valor de salida es categórico mientras que es una

regresión cuando el valor de salida es numérico. La clasificación puede ser binaria o multivariada. Por otro lado el aprendizaje no supervisado incluye la agrupación (cluster), análisis de componente principal y la generación de redes adversarias.

1.3.3. Redes Neuronales Profundas Convolucionales

Las Redes Neuronales Convolucionales en aprendizaje profundo se encuentran en un momento crítico para la visión por computadora para los algoritmos supervisados y no supervisados, estas han roto los records de las tareas tradicionales (Wang y col., 2017). Las partes que componen una CNN son capas convolucionales, pooling y completamente conectadas. La función de las capas convolucionales es la identificación de patrones, líneas y bordes. Cada capa oculta de la CNN consiste de capas convolucionales que convulsionan un arreglo de entrada con kernels convolucionales con pesos parametrizados. Los múltiples kernels generan diversas imágenes de características que pueden lograr diversas tareas de visión por computadora como segmentación y clasificación. Entre las capas convolucionales mapas de características son progresivamente ingresados a las capas de pooling. Las capas de pooling transfieren el valor máximo o promedio y entonces se reduce el tamaño de los mapas de características. Este proceso captura características robustas de las imágenes respecto a la posición y la forma. Empíricamente el proceso de pooling máximo es utilizado por lo general. La arquitectura de la CNN esta compuesta de capas alternadas de convolución y pooling. Para la clasificación y regresión se utilizan las capas completamente conectadas al final de la arquitectura de la CNN y proveen la decisión final. Durante el entrenamiento se produce una pérdida (loss) la cual es estimada por la diferencia entre el valor de la etiqueta y el del resultado, mientras que en las tareas de segmentación las capas de convolución y las de sobre muestreo son unidas al final de las capas de pooling para reconstruir el tamaño de la imagen de entrada. En este caso la pérdida de entrenamiento es evaluada al diferenciar la imagen etiqueta enmascarada y la imagen reconstruida a la salida de la CNN. 1.2

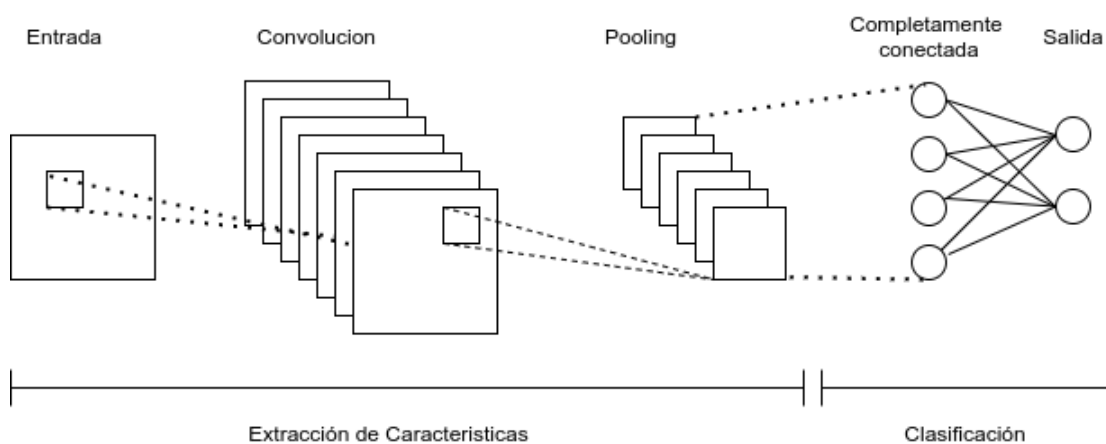


FIGURA 1.2: Red Neuronal Profunda Convolutional

Dado de la que CNN esta compuesta por múltiples capas el numero de parámetros puede alcanzar dimensiones extremadamente grandes (millones). Esto significa que

grandes cantidades de datos son necesarias para el proceso de entrenamiento para lograr una precisión competente. La cantidad de datos que necesitamos dependerá del propósito de la tarea y de las características de la imagen, por ejemplo al menos 1000 imágenes de cada clase son necesarias para obtener un resultado competente en tareas de clasificación. Sin embargo este proceso de adquisición de datos es muy complicado sobretodo si se requiere que las imágenes tenga etiqueta ya que algún experto en el área que se esta investigando debió de realizar esta tarea para que los datos sean confiables. para tratar de ayudar con esta situación se puede utilizar la generación sintética de datos, en el caso de las imágenes esta operación se realiza manipulando la imagen (rotando, trasladando, escalando, invirtiendo, etc) para generar una muestra de similares características que nos permita ampliar el conjunto de datos.

1.3.4. Aplicaciones en Radiología

En esta sección se describen algunas de las tareas mas comunes en las cuales se utiliza el aprendizaje profundo aplicado a imágenes médicas.

Clasificación de Imágenes

Una de las tareas clave para los radiólogos es un diagnostico diferencial apropiado de las imágenes medicas de los pacientes, su tarea de clasificación un gran rango de aplicaciones desde la determinación de la existencia o inexistencia de una enfermedad hasta identificar el tipo de malignidad en el caso de existencia. El uso reciente de las DNN especialmente las CNN han mejorado el desempeño de los clasificadores de imágenes en varias aplicaciones medicas como el diagnostico de tuberculosis, retinopatía diabética, y cáncer de piel (Lakhani y Sundaram, 2017) (Gulshan y col., 2019) (Esteva y col., 2017). Debido a que las imágenes médicas contienen varios tamaños y tipos de patrones de enfermedades complejas, seria complicado para los modelos de CNN entrenar directamente patrones muy complicados de enfermedades. Estos problemas complejos podrían ser resueltos con un una estrategia que involucre un entrenamiento gradual.

El aprendizaje profundo requiere de una gran cantidad de datos para evitar el sobreentrenamiento y mejorar su desempeño, aunque en la practica es difícil alcanzar las dimensiones ideales de imágenes medicas en enfermedades de baja incidencia.

Detección de Objetos

La detección de objetos consiste en encontrar y categorizar objetos, en imágenes biológicas una técnica de detección es identificar áreas donde están localizadas las lesiones del paciente por medio de cajas marcadas con coordenadas. La detección de objetos basada en aprendizaje profundo puede ser compuesta por 2 tipos, uno de ellos son los algoritmos basado en regiones propuestas donde se generan un conjunto de parches usando un algoritmo de búsqueda selectiva en imágenes de entrada, después el modelo entrenado decide cuando múltiples objetos existen en cada parche y clasifica los objetos basado en una región de interés (ROI por sus siglas en ingles) .Específicamente la red de región propuesta fue desarrollada para incrementar la velocidad del proceso de detección.La otra técnica realiza una detección de objetos por medio de métodos de regresión como una red de una fase. Estos enfoques detectan y encuentran directamente

las coordenadas de los bordes de las cajas y las probabilidades de clases de los píxeles de imagen en toda la imagen. Los métodos de una etapa son más rápidos pero los de 2 etapas muestran mejor precisión. Varios algoritmos para la detección de objetos para imágenes médicas son basados en etiquetas por píxel o etiquetas por cajas de borde. Para adquirir estas etiquetas para la detección de patrones de enfermedades es demasiado complicado además de inevitable en entornos médicos. Para sobrellevar el costo de los datos anotados se debe explotar la técnica de transferencia de conocimiento con pesos pre entrenados de modelos de datos mucho más grandes para después realizar un ajuste fino con el conjunto reducido de imágenes. (Redmon y col., 2016)

Segmentación de Imágenes

Las imágenes médicas proveen de mucha información, diversos algoritmos automáticos de segmentación y registro automáticos han sido estudiados y propuestos para su uso en estudios, el aprendizaje profundo ha mostrado un desempeño excelente en varias de estas aplicaciones.

El método clásico de segmentación de imágenes está basado en los filtros de detección de bordes y múltiples algoritmos matemáticos. Para mejorar el desempeño asociado a la segmentación las CNN se han ido introduciendo gradualmente. Se han realizado estudios sobre la segmentación de tumores y otras estructuras del cerebro, pulmones, células biológicas y membranas. (Pereira y col., 2017) Estos enfoques utilizan técnicas de parches en 2 dimensiones con CNN y post procesamiento con aprendizaje máquina clásico. Sin embargo el entrenamiento puede tardar demasiado y dependiendo del número de parches puede que el aprendizaje no sea posible.

Algunas arquitecturas han sido propuestas pero debido a la naturaleza de las CNN las imágenes resultantes carecían de la resolución necesaria, una implementación que combina capas de deconvolución permite a la red producir imágenes de segmentos con la misma resolución que la imagen origen. Recientemente diversas técnicas mejoradas de aprendizaje profundo para realizar la segmentación continúan siendo propuestas. Con el incremento del desempeño de la segmentación estudios han conducido a considerar la incertidumbre de las etiquetas, dado el costo del etiquetado se han propuesto técnicas semi y sin supervisión pero estos aún no han superado el desempeño de los sistemas supervisados.

Generación de Imágenes

Mientras que muchas aplicaciones que utilizan CNN han sido utilizadas en imagenología médica es complicado encontrar conjuntos de datos etiquetados, balanceados y con imágenes en alta definición. Para auxiliar en estos casos existen redes adversarias generativas (GAN por sus siglas en inglés) las cuales se utilizan para la generación sintética de imágenes de RX, CT y ROIS de lesiones específicas.

Las GAN son la combinación de 2 diferentes redes neuronales que pueden generar imágenes sintéticas realistas. (Goodfellow y col., 2020), estas redes fueron introducidas en 2014 y por lo general se utilizan para aumentar la cantidad de imágenes y así obtener un conjunto de datos mejor balanceado para una red neuronal que clasifica o segmenta. Con este conjunto de datos balanceado se vio mejorado el desempeño de los sistemas en comparación con los no balanceados.

Otra aplicación en la que se utilizan las GAN es en la detección de anomalías en imágenes médicas, al tratar de generar las imágenes sintéticas el modelo entrenado puede aprender de las imágenes que no presentan lesiones y entonces detectar cuales de ellas presentan algún tipo de anomalía. Los estudios demostraron el potencial de las GAN sin supervisión.

Transformación de Imágenes

Los estudios recientes de transformación de imágenes se puede dividir en 2 categorías, los que utilizan GAN y los que no. Las redes CNN se pueden utilizar para eliminar el ruido de las imágenes . (Valesia, Fracastoro y Magli, 2019) Jain y Seung mostraron técnicas para eliminar el ruido de las imágenes por medio de su arquitectura CNN, se comparo su desempeño con el método de campo aleatorio de Markov produciendo resultados similares.

El uso de GAN en la transformación de las imágenes se ha utilizado para realizar traslación de las imágenes con correspondencia pixel por pixel, este modelo es llamado red pix2pix, para sobrellevar la limitación que requiere este tipo de correspondencia se diseño la arquitectura cycleGAN (Zhu y col., 2017) .Aunque CycleGAN puede utilizarse para desasociar imágenes, la transferencia de estilo para mas de tres dominios requiere demasiados generadores.

Se han utilizado las GAN para generar rebanadas de CT mas delgadas (Kudo y col., 2019), se basaron en el uso de un parche en tres dimensiones.

Existen diversas transformaciones que se pueden realizar con o sin GAN pero no existe un consenso general sobre cuales ofrecen resultados mas satisfactorios respecto al uso de una u otra.

Capítulo 2

Capítulo 2- Marco Conceptual y Contextual

2.1. Introducción

Los coronavirus son virus comunes en muchas especies de animales, como camellos, vacas, gatos y murciélagos. En raras ocasiones pueden infectar a las personas y luego propagarse entre ellas. El MERS, el SARS-CoV-1 y el SARS-CoV-2, que está causando la pandemia actual de la enfermedad por coronavirus 2019 (COVID-19), pertenecen a esta familia de virus .

Este nuevo virus causa una enfermedad respiratoria potencialmente grave y se descubrió por primera vez en la ciudad de Wuhan, provincia de Hubei, China. A partir de aquí, se han detectado miles de casos de COVID-19 en China y millones en todo el mundo (*Coronavirus disease (COVID-19)*). A finales de enero de 2020 la Organización Mundial de la Salud (OMS) declaró el brote como emergencia de salud pública de importancia internacional. Actualmente es una pandemia de enormes proporciones.

La presencia del nuevo virus SARSCoV-2 ha provocado un impacto social, económico y de salud en la población mexicana y en general a nivel mundial. Es importante comprender las razones por las cuales afecta a diversos sectores socioeconómicos de forma diferente y al momento de la realización de esta investigación aun no es definitivo un tratamiento lo suficientemente confiable y basado en evidencia científica para el tratamiento de pacientes con afectaciones severas causadas por la enfermedad.

El virus se propaga de persona a persona por aerosoles generados al hablar o respirar y, más raramente, mediante manos o fómites por contacto con superficies contaminadas. La estabilidad de SARS-CoV-2 es similar a la del SARS-CoV-1 en los diferentes materiales. El SARS-CoV-2 se detecta hasta 3 tres horas en aerosoles y hasta 2 o 3 días en materiales sólidos, aunque la vida media estimada es de 5 y 8 horas, respectivamente. También se ha encontrado en la sangre y en las heces. Parece poco probable que estas vías tengan un peso relevante en la transmisión. Aunque varias especies animales pueden padecer la enfermedad, hay muy pocos casos descritos de posible transmisión de animales a humanos. (*Coronavirus disease (COVID-19)*)

El número de casos que genera un individuo infectado sin medidas comunitarias de aislamiento se estima entre 2.2 y 2.6 . Esto puede variar según comunidades y es mayor en ambientes cerrados. En algunos centros de mayores y refugios la prevalencia y la tasa de contagiosidad fue superior al 30 %. La transmisibilidad del SARS-CoV-2 disminuye mucho con medidas de salud pública y distanciamiento físico, y no parece mostrar signos de debilitamiento en condiciones cálidas y húmedas, aunque algunos

análisis apuntan una menor capacidad de transmisión a temperaturas mayores . Parece que las nuevas variantes del virus, como la británica, tienen mayor capacidad de difusión que la cepa original. (Azghandi y Kerachian, 2020)

No se conoce con seguridad el grado de contagiosidad en cada etapa de la enfermedad. En pacientes que tienen un curso leve, el pico de la carga viral en muestras nasofaríngeas ocurre durante los primeros 5-6 días tras el inicio de síntomas y prácticamente desaparece después del día 10. En algunos pacientes se detectan virus no viables en cultivos durante más tiempo, a los que se les supone una baja capacidad de contagio. Aunque se necesitan estudios para conocer con más detalle la dinámica de la transmisión, el grado de contagiosidad disminuye progresivamente, más rápido en los casos con evolución favorable, hasta desaparecer . Se ha detectado carga viral en muestras nasales y capacidad de contagio en la fase presintomática (1-2 días) y en pacientes asintomáticos o poco sintomáticos. (Chen y col., 2020)

La respuesta inmunitaria aún no se comprende completamente, aunque confiere inmunidad contra la reinfección, al menos temporalmente .

2.2. Manifestación Clínica

El período de incubación más habitual está entre 4 y 6 días. A los 12 días más del 95 % de los casos sintomáticos han desarrollado ya sus síntomas, aunque podría llegar hasta los 14 días. Estos tiempos se utilizan como margen de seguridad en algunas cuarentenas.

La COVID-19 se manifiesta habitualmente como una infección respiratoria aguda, aunque puede ser asintomática o apenas sintomática. Los síntomas más comunes en pacientes hospitalizados al inicio de la enfermedad fueron fiebre, astenia y tos seca. Tuviron disnea una tercera parte de ellos, generalmente tras 5 días de inicio del cuadro. Con menos frecuencia se presentaron mialgia, dolor de cabeza, mareos, dolor abdominal, diarrea, náuseas y vómitos . Con frecuencia aparece pérdida repentina de olfato.

Los niños tienen una enfermedad con síntomas similares aunque más leves que los adultos, los marcadores inflamatorios elevados fueron menos comunes, la linfocitopenia infrecuente, menos casos de neumonía severa y las muertes han sido extremadamente raras . Se ha descrito algún caso de enfermedad de Kawasaki, particularmente la forma incompleta y atípica, en niños con enfermedad por coronavirus. En mujeres embarazadas infectadas, los resultados fetales y neonatales parecen buenos en la mayoría de los casos, aunque el embarazo puede estar asociado a un mayor riesgo de preeclampsia, parto prematuro, bajo peso al nacer o incluso muerte fetal intrauterina.

En casos leves, la recuperación se produce a las 2 semanas. En casos graves, el tiempo desde el inicio de los síntomas hasta la instauración de síntomas graves como la hipoxemia es de 1 semana. Se pueden recuperar en 3-6 semanas o fallecer en 2-8 semanas. Algunos pacientes refieren la persistencia de algunos síntomas más allá de ese periodo. En estos casos se habla de COVID persistente, una entidad todavía poco definida . A los 6 meses, el 76 por ciento de todos los pacientes presentaban al menos uno de los siguientes síntomas: fatiga o debilidad muscular (63 %), dificultad para dormir (26 %), caída del pelo (22 %), dificultad para olfato y gusto (11 y 9 por ciento , respectivamente) y problemas de movilidad (7 %). Se han descrito casos de reinfección, aunque faltan estudios para conocerla con mayor precisión .(*Coronavirus disease (COVID-19)*)

Inicialmente la tasa de pacientes graves o muy graves se acercó al 20 % con una tasa de letalidad del 2,3 %. La tasa de letalidad publicada varía mucho según el país. En

Corea del Sur es del 0,9 %; sin embargo, en Italia y España superó inicialmente el 7 %, sustancialmente más alta que en China y otros países.

La mayoría de pacientes graves tuvieron una edad superior a la media y mayor comorbilidad. La edad, especialmente después de los 60 años, es el factor de riesgo más importante de enfermedad grave. También lo son el sexo masculino y la existencia de comorbilidades (cardiopatía isquémica, fibrilación auricular, ictus, EPOC, hipertensión, diabetes, demencia, cáncer activo en los 5 años previos, hepatopatía crónica, enfermedad renal crónica) (. Además de tener más edad y más comorbilidades, los pacientes que desarrollaron neumonía fueron más propensos a presentar fiebre, mialgia, astenia, disnea, dolor de cabeza, náuseas y vómitos . Se necesitan criterios clínicos y biomarcadores más precisos que puedan ayudar a diferenciar a las personas con más probabilidades de progresar a enfermedad grave . Parece que algunas variantes pueden tener una mortalidad hasta un 61 % mayor que la cepa original.

2.3. Pruebas Diagnosticas

En general, la decisión de realizar una prueba debe basarse en factores clínicos o epidemiológicos y vincularse a una evaluación de la probabilidad de infección con protocolos adaptados a la situación local. Si el manejo del caso lo requiere, deben realizarse pruebas para detectar otros patógenos respiratorios, sin retrasar las pruebas para COVID-19, que deben realizarse siempre, aunque se encuentre otro patógeno respiratorio que pudiera justificar la sintomatología .(*Coronavirus disease (COVID-19)*)

Existen dos tipos de pruebas de utilidad diagnóstica: las basadas en la detección del virus (ARN o antígeno viral) y las basadas en la detección de anticuerpos (IgM o IgG) frente al virus.

La PCR (reacción en cadena de la polimerasa), que se basa en la detección de material genético (ARN) del virus, es la prueba más usada para diagnosticar la COVID-19. Se realiza habitualmente sobre una muestra nasofaríngea, necesita laboratorios especialmente equipados, puede detectar casos aún en fase asintomática y alcanza su punto máximo de utilidad dentro de la primera semana de aparición de los síntomas. Esta prueba puede superar el 20 % de falsos negativos, por lo que debe repetirse si se mantiene la sospecha clínica. Parece que el cultivo del virus habitualmente no es posible más allá del día 8 tras el inicio de los síntomas, lo que se correlaciona con la disminución de la infectividad tras la primera semana . Hay además varias técnicas que detectan el antígeno viral, test rápidos que no precisan de personal experto, aunque son menos sensibles. Estas pruebas de antígenos son cada vez más precisas y de gran utilidad cuando no se dispone de PCR o se necesita un diagnóstico rápido, especialmente en la primera semana, en la que las cargas virales son más altas .

La prueba serológica de anticuerpos busca la presencia de proteínas específicas producidas en respuesta a la infección (IgM e IgG). Los anticuerpos detectados por esta prueba indican que una persona tuvo una respuesta inmunitaria al SARS-CoV-2, aunque la infección fuera asintomática. Es más económica y sencilla que la PCR, tiene gran utilidad epidemiológica y puede ayudar a identificar a los pacientes que tuvieron una infección por el SARS-CoV-2 3 o 4 semanas antes o más.

El marcador serológico más sensible y más temprano son los anticuerpos totales, cuyos niveles comienzan a aumentar a partir de la segunda semana del inicio de los síntomas, aunque con ELISA IgM e IgG pueden ser positivas incluso al cuarto día después

del inicio de los síntomas. Los niveles más altos se producen en la segunda y tercera semana de enfermedad .

La seroconversión de IgM e IgG se produce en todos los pacientes entre la tercera y cuarta semana tras el inicio de la clínica. La IgM comienza a disminuir y alcanza niveles más bajos en la semana 5 y casi desaparece en la semana 7, mientras que la IgG persiste más allá de las 7 semanas . La eficacia de detección por IgM es mayor que la de PCR después de 5 días de inicio de síntomas. La tasa de detección positiva aumenta significativamente (98,6 %) cuando se combina la IgM con PCR para cada paciente en comparación con una sola prueba. (*Coronavirus disease (COVID-19)*)

Se está desarrollando un gran número de vacunas contra el SARS-CoV-2 de diferentes tipos, algunas ya en fase de autorización condicional, con una eficacia superior al 95 % para la enfermedad sintomática y grave . Al mismo tiempo aparecen nuevas variantes para las que las vacunas actuales podrían no ser tan eficaces, aunque hasta este momento, al menos, sí lo parecen.

La rapidez del desarrollo de las vacunas se debe en parte al descubrimiento de las que utilizan el ARN mensajero (ARNm). Este ARNm tiene instrucciones para producir una proteína de superficie que el virus SARS-CoV-2 utiliza para entrar en las células. Tras la inyección de la vacuna, algunas células leen las instrucciones del ARNm y producen temporalmente la proteína de pico. El sistema inmunitario de la persona receptora, tras reconocer esta proteína como extraña, produce anticuerpos y activa los linfocitos T, lo que la protegerá cuando entre en contacto con el virus SARS-CoV-2. La partícula de ARNm es posteriormente degradada y eliminada por enzimas. En este proceso no está involucrado ningún virus vivo ni material genético que altere el núcleo de las células.

Teniendo en cuenta los estudios y seguimiento realizados hasta la fecha, estas vacunas son muy seguras. Entre los efectos adversos conocidos destaca por su gravedad un aumento de las tasas de eventos tromboembólicos venosos, incluida la trombosis venosa cerebral, por el desarrollo del síndrome de trombocitopenia inmunitaria protrombótica inducida por vacuna (VIPIT), entre los receptores de Ad26.COV2.S y especialmente ChAdOx1-S (AstraZeneca). Sin embargo, los riesgos absolutos de estos eventos fueron pequeños y los hallazgos deben interpretarse a la luz de los efectos beneficiosos probados de la vacuna . Por otro lado, con las vacunas basadas en ARNm se han declarado casos muy raros de anafilaxia, entre 2.5 y 4.7 casos por millón .

2.4. Tratamiento de la Enfermedad

No hay evidencia actual de ensayos clínicos para recomendar una terapia única y eficaz para pacientes con infección sospechada o confirmada de COVID-19. Solo algunas opciones tienen utilidad limitada a situaciones clínicas concretas . Lo más importante es asegurar un tratamiento de soporte precoz, administrar antimicrobianos para tratar los posibles agentes etiológicos del síndrome de distrés respiratorio agudo y adaptar el tratamiento a las condiciones de cada persona y sus comorbilidades .

Aunque muchos medicamentos tienen actividad in vitro contra diferentes coronavirus, actualmente no hay evidencia clínica sólida que respalde la eficacia y seguridad de ninguno contra cualquier coronavirus en humanos, incluido el SARS-CoV-2. (*Coronavirus disease (COVID-19)*)

El remdesivir parece superior a placebo en acortar tiempo de recuperación, pero no reduce la mortalidad ni la necesidad de ventilación mecánica . Un análisis de subgrupos

indicó que el remdesivir posiblemente aumente la mortalidad en los enfermos críticos y reduzca la mortalidad en los que no están gravemente enfermos (*Coronavirus disease (COVID-19)*). La EMA autorizó su uso en pacientes con COVID-19 con neumonía que requieren oxígeno suplementario.

El tocilizumab, un anti-receptor de la interleucina-6, puede ser útil en combinación con dexametasona en algunos pacientes hospitalizados que presentan una descompensación respiratoria rápida. Reduce la mortalidad por todas las causas en el día 28 en comparación con la atención estándar sola, aunque probablemente tenga poco o ningún efecto sobre la mejoría clínica.

Los estudios publicados no respaldan el uso rutinario de corticosteroides en COVID-19, pero algunos hallazgos sugieren que la metilprednisolona podría reducir la tasa de mortalidad en las formas más graves de la afección, por lo que estarían indicados en pacientes gravemente enfermos que reciben oxígeno suplementario o ventilación mecánica.

La administración de corticosteroides sistémicos se asocia a una menor mortalidad por todas las causas a los 28 días en pacientes críticamente enfermos con COVID-19.

Todos los adultos hospitalizados con COVID-19 deben recibir trombopprofilaxis farmacológica con heparinas de bajo peso molecular, a menos que el riesgo de hemorragia supere el riesgo de trombosis. En el contexto de la trombocitopenia inducida por heparina, se recomienda fondaparinux.

No existen evidencias claras de que lopinavir-ritonavir disminuyan el tiempo para la mejoría clínica o la necesidad de ventilación mecánica ni reduzcan la mortalidad (*Coronavirus disease (COVID-19)*). Además, su potencial hepatotoxicidad es un factor relevante, ya que la COVID-19 también puede alterar la función hepática.

El tratamiento con plasma convaleciente, hidroxiclороquina o vitamina D en pacientes con COVID-19 no se asoció a una mayor supervivencia ni otros resultados clínicos positivos.

La rehabilitación respiratoria de 6 semanas puede mejorar la función respiratoria, la calidad de vida y la ansiedad de los pacientes ancianos con COVID-19, pero tiene utilidad en la depresión en los ancianos. (*Coronavirus disease (COVID-19)*)

2.5. Perfil de Una afectación en los Pulmones Causada por COVID-19

Una parte esencial en los algoritmos de reconocimiento de patrones son las características que definen una repetición de un suceso. En este caso se trabajara con dos diferentes técnicas de obtención de imágenes (CT y XR), esto implica que se requiere conocer las representaciones visuales de la problemática causada en específico por el COVID-19 a nivel de los pulmones. Existen referencias bibliográficas de especialistas que detallan cuales son los hallazgos relevantes en el diagnostico visual a partir de estas imágenes.

El análisis de estas evidencias es de suma importancia, se puede realizar un pre procesamiento de la imagen para facilitar y mejorar el proceso de reconocimiento de características asociadas a COVID-19 en afecciones pulmonares. Por lo general los algoritmos de Aprendizaje Profundo para imágenes requieren que estas estén normalizadas,

por lo que una aplicación de filtros deberá de ser utilizado para que las imágenes tengan niveles de brillo y contraste similares.

Es importante señalar que no todas las personas enfermas con el virus SARS-COV-2 presentan afecciones pulmonares, estas existen, por lo general, en personas mayores, con comorbilidades o con el sistema inmune deprimido.

En las siguientes secciones se describen las características observables en cada una de las técnicas de adquisición de imágenes que se utilizaran en esta investigación.

Perfil de CT de Torax en Pacientes con COVID-19

En un estudio clínico sobre las características en imágenes CT de tórax de pacientes con COVID-19 (Chung y col., 2020) se notaron las siguientes evidencias:

- las opacidades de tipo vidrio deslustrado se presentan por lo general en personas afectadas.

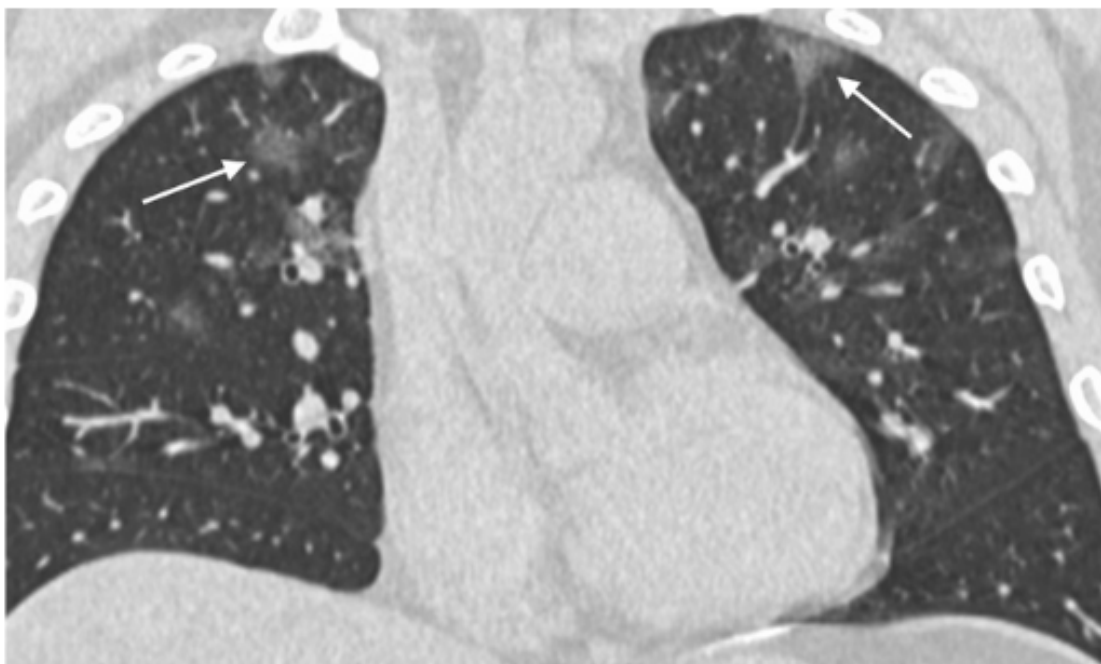


FIGURA 2.1: Imagen CT de Tórax Coronal. Se muestra las opacidades de tipo vidrio deslustrado con morfología redonda.

- Aproximadamente la mitad de ellos presentaron afectaciones en 5 lóbulos y solo el 5 por ciento de los pacientes no presentaron ninguna afectación en los lóbulos pulmonares.
- La mayoría de los pacientes con opacidades las presentaron en ambos pulmones.
- el 33 por ciento de los pacientes presentaron opacidades y/o vidrio deslustrado con morfología circular 2.1 , 14 por ciento presentaron principalmente anomalías lineales y 19 por ciento un patrón de empedrado 2.2 , un 21 por ciento distribución periférica 2.3.

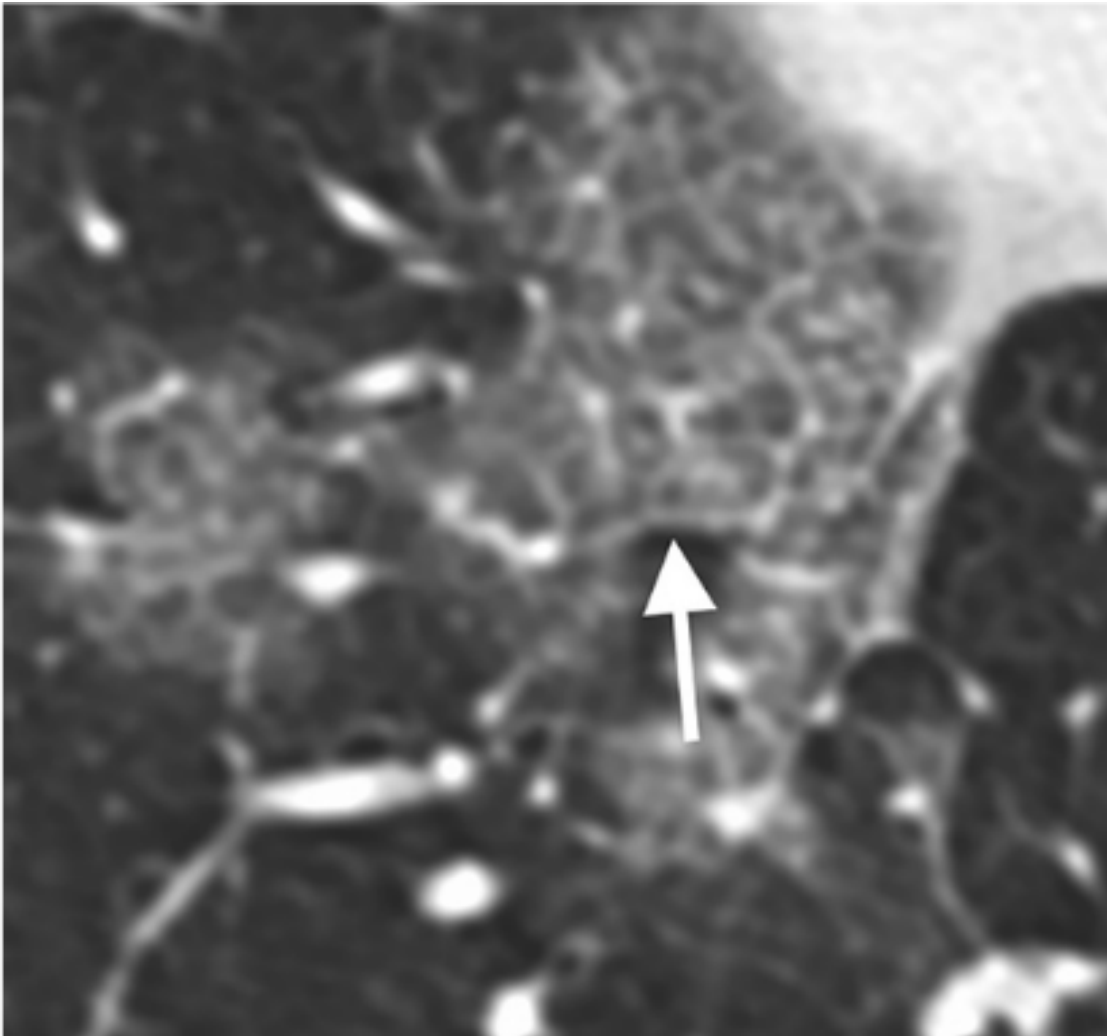


FIGURA 2.2

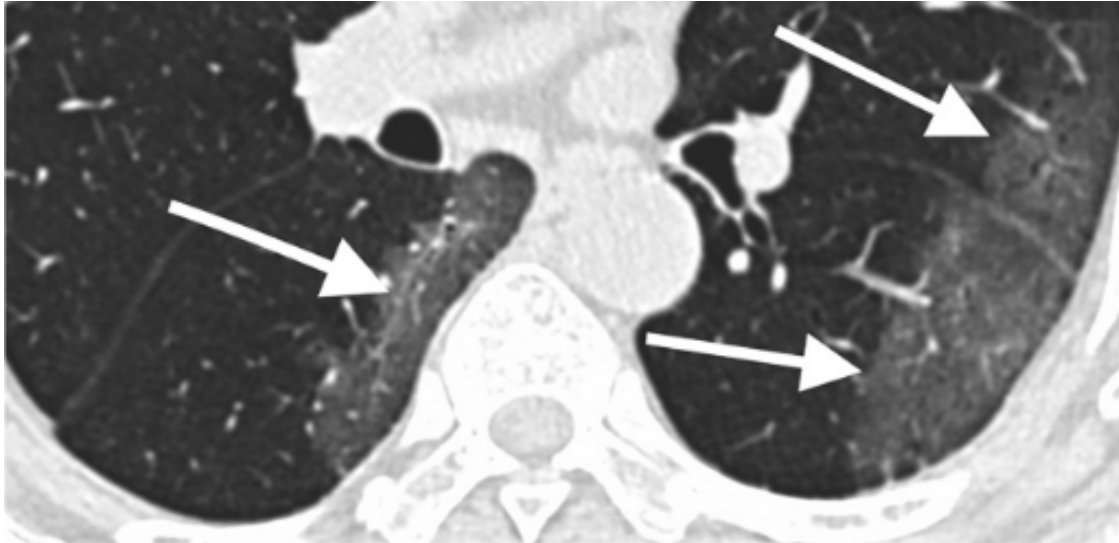


FIGURA 2.3

- Ningún paciente tuvo cavitaciones en el pulmón, nódulos pulmonares discretos, efusiones pleurales, linfadenopatía, enfisema pulmonar o fibrosis.

Por lo general las imágenes mas utilizadas para diagnostico con CT son las perspectivas coronales.

Perfil de XR de Torax en Pacientes con COVID-19

Las imágenes de tórax muestran las alteraciones pulmonares posteriores a las etapas tempranas de la enfermedad. Además las imágenes XR tiene una menor nitidez que las CT.

Las alteraciones que se presentaron en los pacientes con COVID-19 fueron: consolidaciones en el pulmón en un 57% 2.4, vidrio deslustrado en un 62% , nódulos en un 23% y opacidades nodulares-reticulares en un 66%. La mayoría de los pacientes (69%) tuvieron afectación bilateral. Los pacientes mas afectados fueron aquellos con una edad superior a los 60 años.

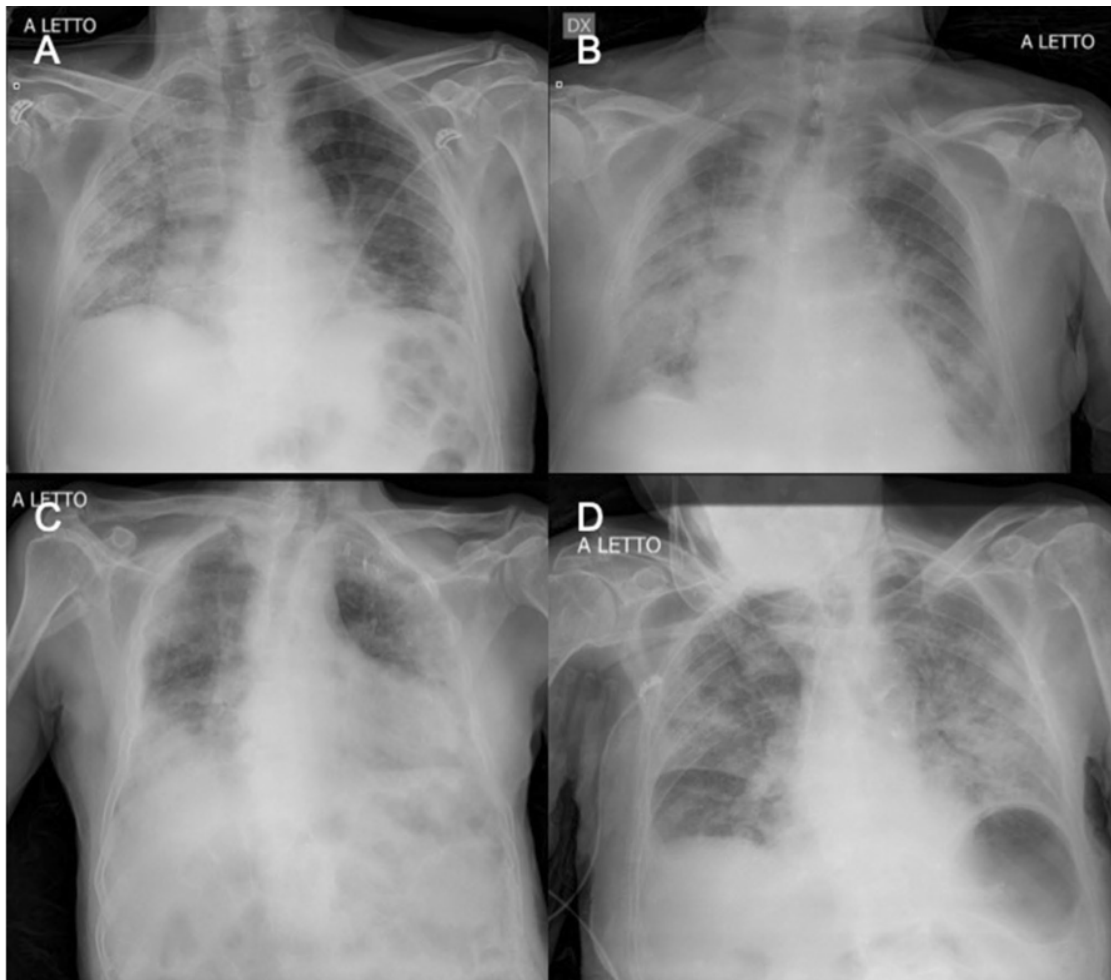


FIGURA 2.4: Paciente con daño avanzado en los pulmones y consolidación bilateral.

Capítulo 3

Capitulo 3 - Marco Metodológico

3.1. Conjunto de Datos

3.1.1. introducción

Una parte esencial para el uso de cualquier algoritmo de proyección, segmentación, agrupación o clasificación de información es la capacidad de tener datos confiables que representen fielmente la población que se va a tratar.

En específico para las redes neuronales artificiales de aprendizaje profundo los datos requeridos para entrenar los modelos de redes necesitan ser muy grandes, entonces la pregunta que surge inicialmente es ¿ Qué tan grande debe ser el conjunto de datos ?. Es una respuesta complicada de responder, algunos estudios muestran que por regla general al utilizar aprendizaje profundo bastaría con tener al menos 1000 imágenes por cada clase, pero esto nos daría resultados un poco tendenciosos por lo que cuantos mas datos tengamos mejor sera el desempeño de nuestra red neuronal, se menciona que esta proporción se incrementa de manera logarítmica respecto al incremento en los datos que se utilicen al momento de entrenar. Todo esto tiene un limite y no necesariamente por tener un conjunto de datos enorme el desempeño mejorara infinitamente como lo muestra Joulin y col., 2015 donde se utilizaron 100 millones de imágenes, el autor reporto que a partir de 50 millones de datos el desempeño de la red convolucional descendió un poco, también debido a que se robusteció la generalización de esta.

Otro punto esencial dentro del conjunto de datos es el desbalance de estos, debido a que se trata e una enfermedad emergente el conjunto de imágenes disponibles es limitado, al momento de realizar este trabajo, existe una cantidad basta de imágenes de tórax tanto tomografías como rayos X de pacientes normales y con diversas enfermedades como neumonías provocadas por diversas enfermedades o cáncer por mencionar algunas. Es importante tomar en cuenta este aspecto para que el entrenamiento de nuestra red sea lo mas eficiente posible, tratar de satisfacer los requerimientos de datos que la red necesite sin saturar innecesariamente una clase en específico.

Como parte de un estudio que cuente con resultados robustos se realiza un análisis de los conjuntos de datos que cuentan con validación por parte de personal médico, algunos de ellos son utilizados en proyectos similares al que se realiza y una parte que se considera conveniente es que en cada uno de ellos existe una diversidad de imágenes, a lo que se refiere con diversidad es que estas han sido obtenidas por diferentes dispositivos (tomógrafos y equipos de rayos X) en el caso de las tomografías estas se han capturado con diversos grosores y resoluciones, por su parte los rayos X se capturan con diferentes metodologías como PA(Posterior Anterior) y AP (Anterior Posterior) así

como con diferentes potencias de kilo-voltaje lo que entrega imágenes con diversa penetración de rayos X, en consecuencia imágenes con características diferentes, también con diferentes resoluciones. Esta búsqueda y recopilación de información tiene como fin que el sistema sea entrenado con una diversidad lo suficientemente buena como para que a posterior se pueda clasificar con una precisión adecuada cualquier imagen sin importar con que equipo haya sido capturada.

3.1.2. Datos Públicos de Imágenes de Tórax de Pacientes con COVID-19

El área de interés de este trabajo es el estudio de imágenes de pacientes con COVID-19 que presenten síntomas de afectaciones en los pulmones, ya con anterioridad se ha especificado cuales son las características que presentan estos ante la enfermedad. Para que los resultados que se logren obtener tengan cierta validez se requiere que estos datos sean fiables y se encuentren bien documentados por personal médico. Los repositorios públicos al inicio de la pandemia eran muy escasos y diversos proyectos utilizaron estos pocos datos para realizar las primeras redes clasificadoras de las imágenes de tórax de pacientes.

En los inicios de 2021 los conjuntos de datos se consolidaron y ya se encuentran disponibles imágenes de diversas regiones del mundo, etiquetadas por equipos médicos reconocidos, en diversos formatos por ejemplo NII, DICOM, JPG y PNG, algunas de ellas cuentan con el estudio clínico del paciente.

Para este trabajo se utilizan los siguientes conjuntos de datos :

- CheXpert X-Ray Dataset
- covid-chestxray-dataset COVID-19 image data collection
- COVID-CT-Dataset
- Padchest-xray-dataset
- Actualmed COVID-19 Chest X-ray Dataset Initiative
- COVID-CTset
- COVID-19 RADIOGRAPHY DATABASE

En las siguientes secciones se detalla las características de cada uno de ellos.

CheXpert X-Ray Dataset

Es un conjunto de datos desarrollado para la interpretación de radiografías de tórax, que incluye etiquetado por parte de radiólogos. El conjunto de datos es desarrollado para aplicaciones en las que se desea que sistemas automáticos de detección de enfermedades pulmonares puedan auxiliar a personal médico a sus diagnósticos. Es creado por el grupo de aprendizaje máquina de la universidad de Stanford. Consta de 224,326 radiografías de tórax de 65,240 pacientes. El etiquetado de este conjunto de datos se realizó basándose en el glosario de la asociación de Fleischner, la cual es la máxima autoridad en la investigación e interpretación de imágenes radiológicas. Dentro de el se

COVID-CT-Dataset

Es un conjunto de datos validado por un radiólogo experimentado del hospital de Tongji en Wuhan China. Cuenta con 349 imágenes tomográficas de tórax de 216 pacientes con COVID-19. Las imágenes provienen de diversos documentos e investigaciones de medRxiv, bioRxiv, NEJM, JAMA, Lancet, etc. Se ha conservado la resolución y tamaño de la imagen tratando de conservar sus características originales. Se preservan los meta datos como por ejemplo el diagnóstico del paciente, edad, ubicación y reporte radiológico.

Contiene 349 imágenes de 216 pacientes con COVID-19 y 463 de pacientes sin la enfermedad. (Zhao y col., 2020)

Padchest-xray-dataset

Es un conjunto de datos que contiene imágenes de rayos x en alta resolución de al menos 160,000 imágenes de 67,00 pacientes, obtenidas del hospital de San Juan en España, fueron clasificadas en 19 diagnósticos, el 27 por ciento de ellas fueron etiquetadas manualmente el resto se realizó en un proceso automático por medio de una red neuronal artificial recurrente. Cuenta con más de 10,000 imágenes de pacientes con manifestaciones de neumonía.

Actualmed COVID-19 Chest X-ray Dataset Initiative

Es un conjunto de datos generado por un grupo de investigadores de diversas universidades e institutos entre ellos la universidad Waterloo de Canadá, el National Research Council Canada y el Hospital de Selayang. Fue desarrollado con el fin de mejorar los modelos de detección de COVID-19 por medio de imágenes de tórax.

COVID-CTset

Es un conjunto de datos que contiene series tomográficas de 377 personas. Hay 15589 imágenes de personas con la enfermedad COVID-19 y 48269 de personas sanas. Fue recopilado en el centro médico de Negin ubicado en Sari, Iran. Las series son capturadas en formato DICOM en una resolución de 512 x 512 píxeles. Este conjunto de datos cuenta con la información de los pacientes como sexo, edad, padecimientos, diagnóstico médico, etc. (Rahimzadeh, Attar y Sakhaei, 2020)

Las imágenes nativamente están en tiff por lo que son convertidas a formato JPG sin compresión para evitar la pérdida de información, también requieren de un procesamiento adicional ya que fueron procesadas en un tono de grises de 16 bits por lo que a la vista aparecen oscuras y al mezclarse con los otros conjuntos de datos pueden provocar un entrenamiento y clasificación errónea.

COVID-19 RADIOGRAPHY DATABASE

Un conjunto de investigadores de la universidad de Qatar, Doha y la universidad de Dhaka en Bangladesh juntos con médicos de Pakistán y Malasia crearon este conjunto de datos de Rayos X de Tórax de pacientes positivos a COVID-19, normales y con

neumonía viral. Contiene 219 pacientes COVID-19 , 1341 Normales y 1345 con neumonía viral para dar un total de 3616 imágenes de COVID-19 , 10,192 normales y 7565 de neumonía viral .

3.1.3. Creación de Conjunto de Datos Propio

Una parte fundamental de esta investigación es tener una representación de la enfermedad en la población Mexicana, por lo tanto es muy conveniente tener imágenes de personas que hayan presentado afectaciones en los pulmones y que además residan y hayan nacido en México. Con el fin de cubrir esta condición que consideramos de alta importancia se logro obtener del "Hospital Puebla" ubicado en la ciudad de Puebla imágenes de tórax de personas con afectaciones pulmonares causadas por COVID-19 que fueron tomadas con un tomógrafo computarizado. Se logro obtener un total de 185 estudios con un promedio de 700 por serie de cada paciente.

Características de las Imágenes

Las imágenes están capturadas en una resolución de 512 x 512 píxeles, con una profundidad en escala de grises de 16 bits. El grosor de cada slice (rebanada) es de 2.5 mm. Estas características hacen posible la generación de un conjunto de datos robusto y de alta calidad. Para conservar la confidencialidad de los pacientes se eliminaron los datos sensibles, este conjunto de datos incluye información como sexo, edad, padecimientos por mencionar algunos pero por el alcance que se tiene en el proyecto fueron eliminados y no son tomados en cuenta, en trabajos posteriores pueden ser utilizados para realizar diversos tipos de clasificación y diagnóstico de la enfermedad.

Procesamiento de imágenes en formato DICOM

Una tomografía computarizada (TC) comprende muchas decenas, si no cientos, de exposiciones separadas de rayos X que son adquiridas por el escáner CT y compuestas para formar una imagen cohesiva del sujeto de la tomografía computarizada.

Con todas estas imágenes compuestas que se toman y combinan, es necesario que haya una manera de separar al paciente en datos significativos que puedan ser interpretados por médicos y radiólogos capacitados.

Inicialmente estas imágenes están almacenadas en un formato DICOM (Data Interchange Standard for Biomedical Imaging) el cual es el estándar médico para diversos tipos de estudios, dentro de el, además de encontrarse las imágenes o vídeos se puede almacenar información sobre el paciente y su diagnóstico además de que se incluye los métodos de obtención de las imágenes y características de los equipos utilizados para este propósito. Una imagen en el formato DICOM es almacenada en un arreglo de píxeles con metadatos asociados. Los metadatos pueden incluir información acerca de la imagen, como el dispositivo que se utilizó para capturar la imagen así como algunos parámetros de esta. Mientras que algunos datos son obligatorios existen algunos que son opcionales por lo que generalmente los archivos DICOM no contienen siempre la misma cantidad de metadatos.

Para la utilización de éstas imágenes se tuvieron que implementar mecanismos para la adquisición, preparación y procesamiento para que fuera posible su uso dentro del modelo de red neuronal que se propone. En primera instancia se adquieren las series

Elemento	Unidades Hounsfield
Aire	-1000
Pulmón	-500
Grasa	-100 a -50
Agua	0
Sangre	+30 a + 70
Músculo	+10 a +40
Hígado	+40 a +60
Hueso	+700(esponjoso) a +3000(cortical)

CUADRO 3.1: Unidades Hounsfield

completas y se asignan a un formato de variable que se pueda manipular en este caso en arreglo de numpy, posteriormente se realiza una normalización de acuerdo a las características que se desean maximizar.

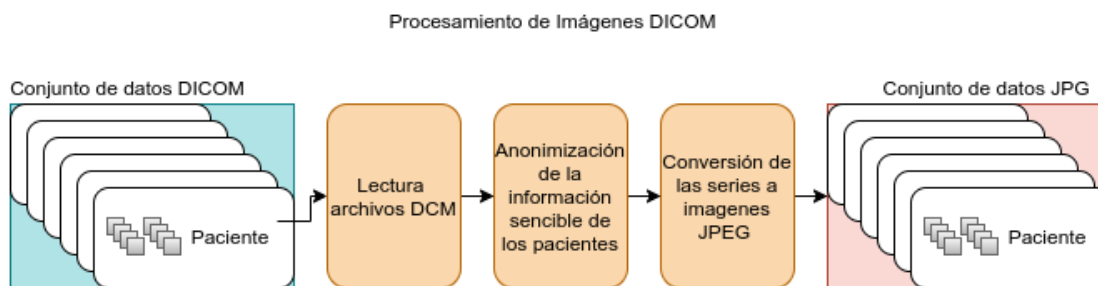


FIGURA 3.2: Adquisición de imágenes de archivos DICOM

Los diferentes sistemas y estructuras anatómicas del cuerpo tienen diferentes niveles de densidad (piense en la grasa de baja densidad en comparación con el hueso de alta densidad) y, por lo tanto, la capacidad de identificar la densidad variable a partir de los rayos X de una manera estandarizada anunció la invención de la TC. escáner.

Para determinar la densidad tisular diversa, Sir Godfrey Hounsfield, entre otros, pudo medir el valor de atenuación (reducción en la fuerza de un haz de rayos X) causado por la absorción de diferentes estructuras anatómicas. El valor de esta atenuación se convirtió en la Unidad Hounsfield (HU). Con una línea de base de 0, siendo la densidad del agua pura, la escala de Hounsfield puede ser tanto negativa como positiva. Cuanto mayor sea el valor, más densa será la materia. 3.1

Una vez calculadas, las imágenes de radiología en los planos axial, coronal y sagital se colorean en escala de grises según su valor HU. La densidad ósea es muy alta y, por lo tanto, es más blanca que los tejidos blandos de baja densidad como la piel, la grasa e incluso el aire de los pulmones ???. Por este motivo se realiza una conversión HU(Hounsfield Units) para obtener una mejor percepción del área del pulmón.

En el histograma de las imágenes con la referencia de HU podemos notar que existe un extremo el cual no aporta información relevante para el estudio que estamos realizando por lo cual después de la conversión y normalización se puede utilizar la imagen con una mejor apreciación del área pulmonar. Claramente en la figura 3.3 podemos notar la existencia de un tono gris en el área corpuscular esta característica es inherente a

la adquisición de las tomografías y radiografías, dependiendo del área que se este estudiando requerimos de herramientas que logren resaltar los elementos de estudio que necesitamos observar de manera mas detallada o con un mayor contraste. 3.4

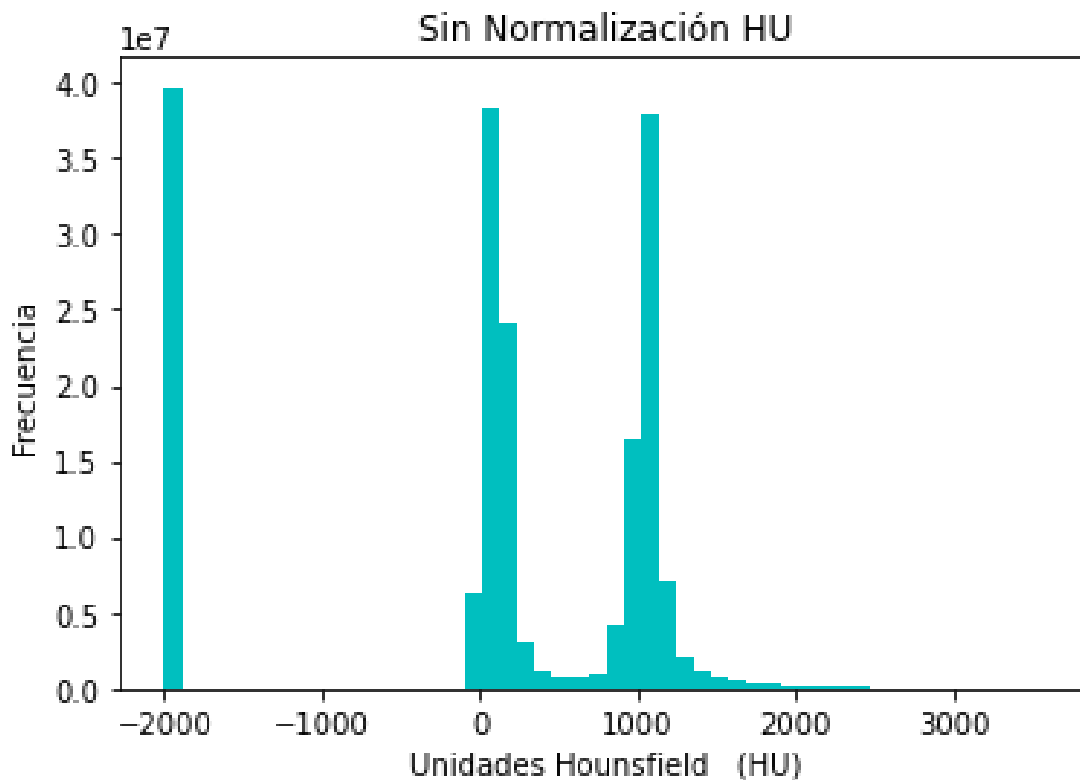


FIGURA 3.3: Histograma de imagen sin HU

En un despliegado de las imágenes del plano axial podemos notar la diferencia de las frecuencias de HU. 3.5 3.6

3.1.4. Conjunto de Datos Final

De cada uno de los conjuntos de datos se trato de utilizar las imágenes que mejor cumplieran con los requerimientos del proyecto, esto es, tener una buena calidad, tener fundamento en un diagnostico por PCR o con validez en el área médica, tener una buena resolución, se trato de tener control sobre la cantidad de imágenes por paciente ya que en algunos conjuntos de datos públicos la cantidad de imágenes era muy grande pero los pacientes eran reducidos en esos casos solo se tomaron alrededor de 100 imágenes por paciente para tener un conjunto de datos mejor distribuido y mas robusto con el cual el modelo de red neuronal pueda generalizar de mejor forma. Se logro obtener un conjunto de datos lo suficientemente grande para realizar los experimentos que requieren un numero masivo de muestras como lo son las redes de aprendizaje profundo. 3.2

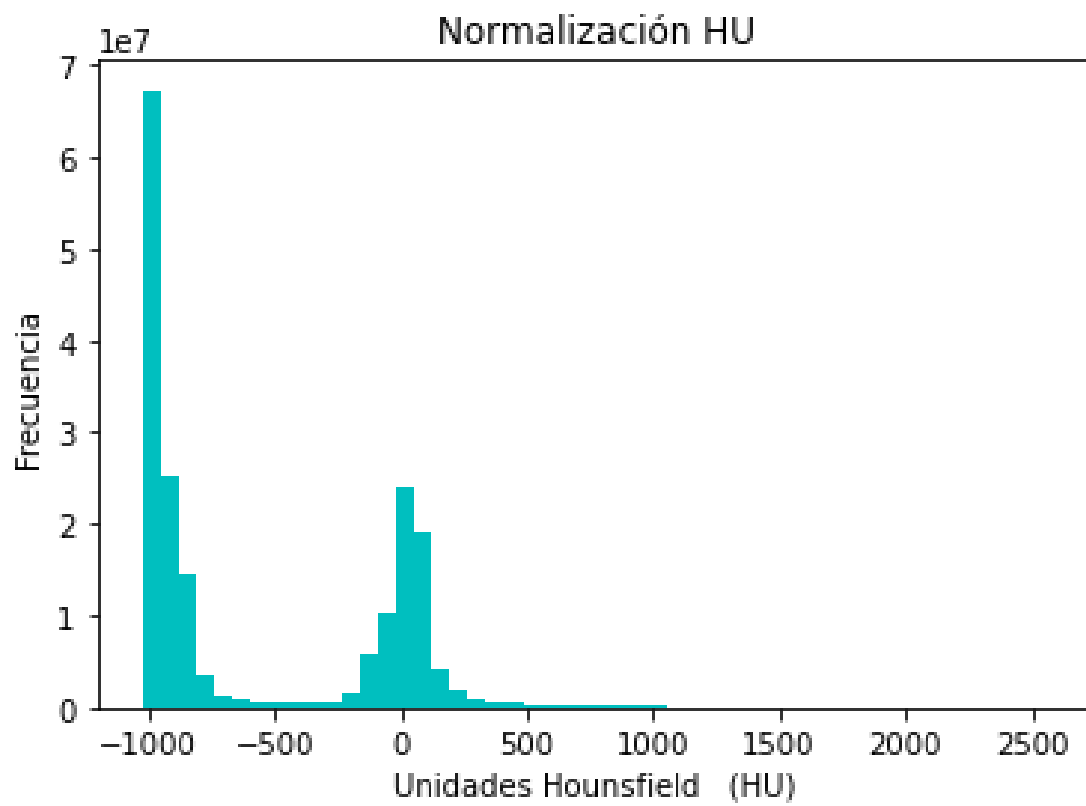


FIGURA 3.4: Histograma de imagen con HU

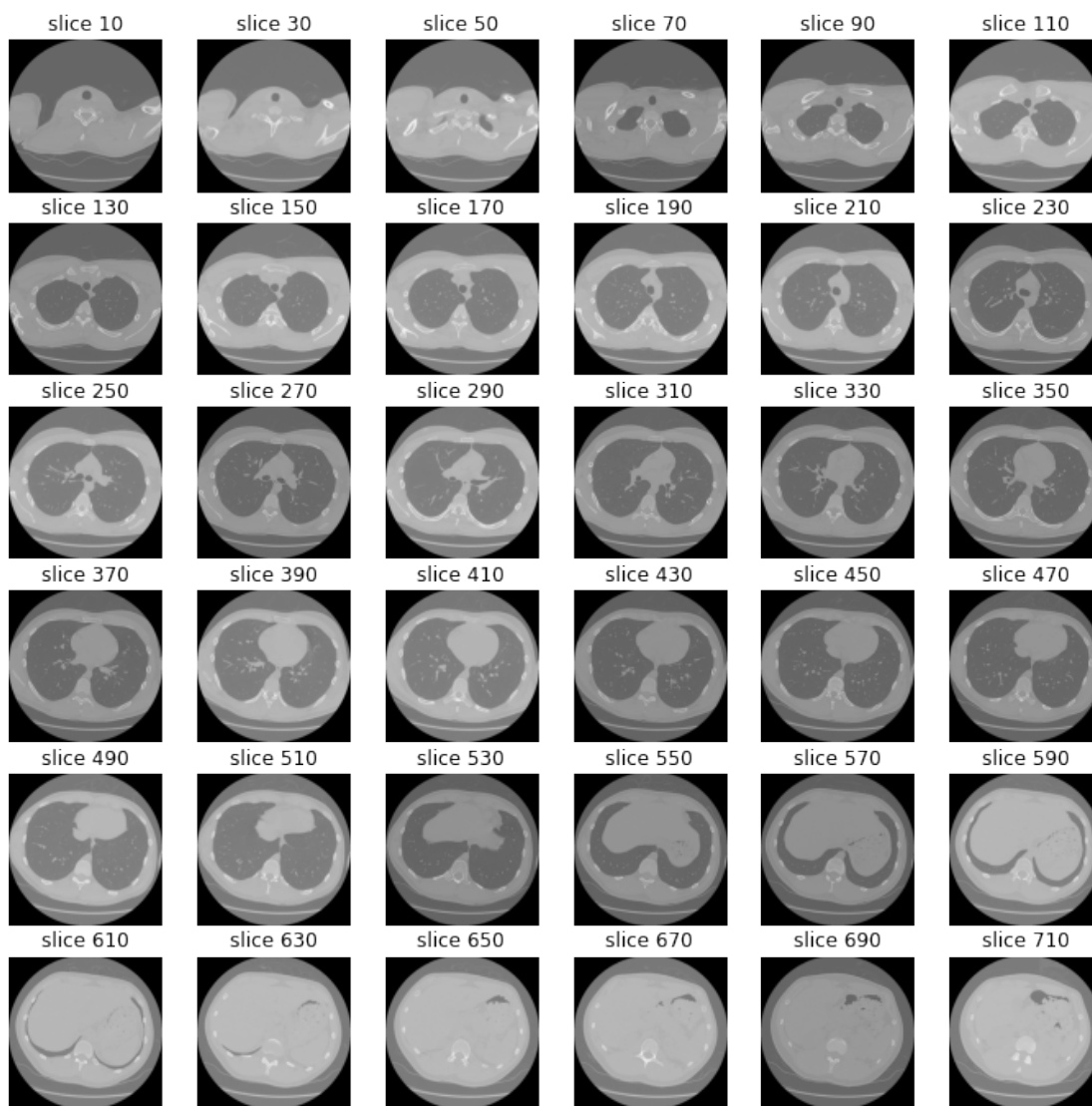


FIGURA 3.5: Planos Axiales Sin Normalización

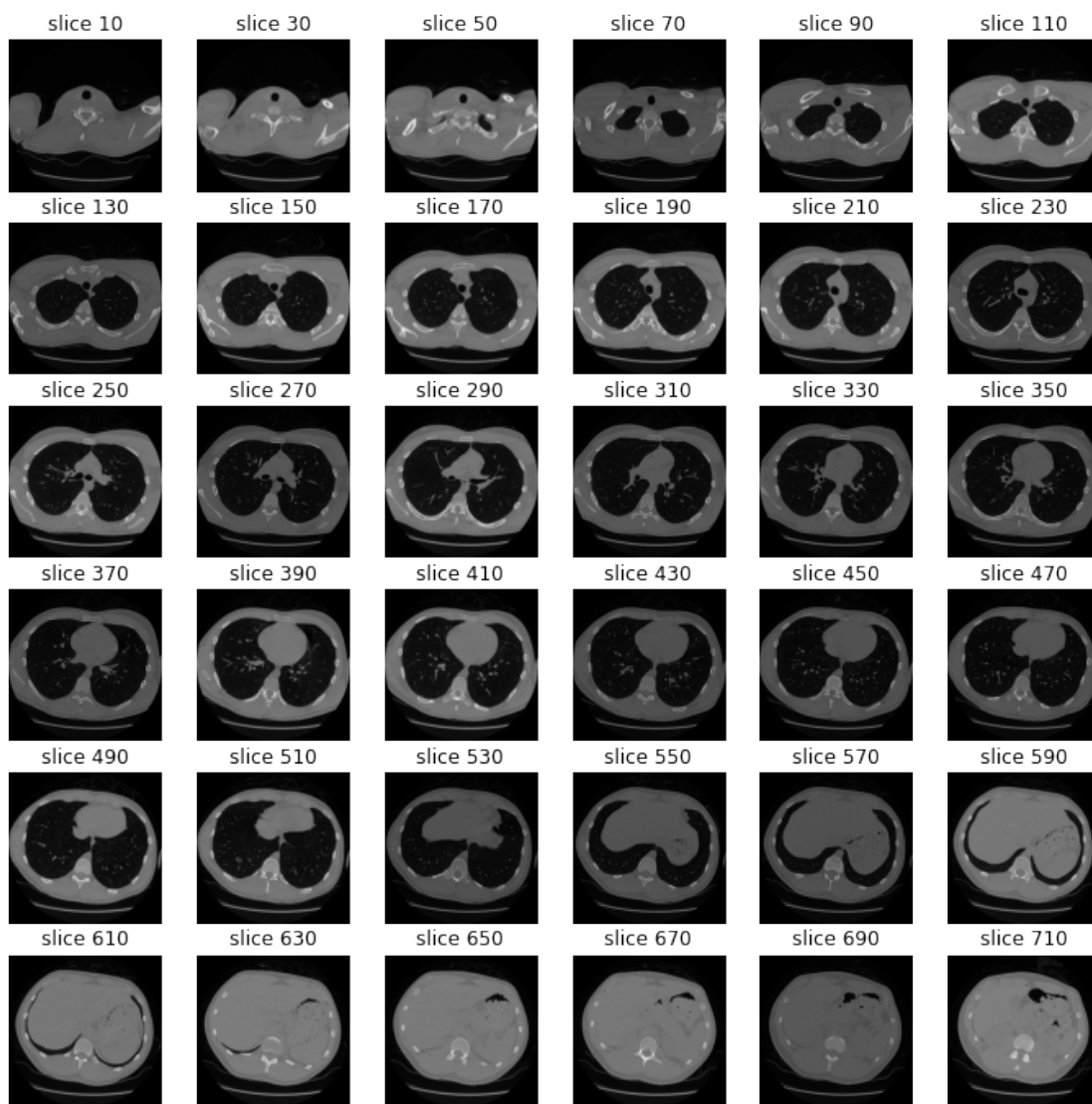


FIGURA 3.6: Planos Axiales Con Normalización

Imágenes de Torax			
	Neumonía COVID-19	Neumonía No COVID-19	Normal
CT	70,180	60,536	66,450
Rayos X	35,890	40,506	34,670

CUADRO 3.2: Conjunto de Datos para entrenamiento de la Red.

3.1.5. Preprocesamiento de las Imágenes Médicas

Antes de que las imágenes puedan ser utilizadas para alimentar un algoritmo de inteligencia artificial realizamos un conjunto de operaciones sobre ellas para lograr obtener los mejores resultados y que las características de los patrones reconocidos en los perfiles de la enfermedad tanto en rayos X como en tomografías computarizadas sean aprendidos con mayor facilidad para la red. Estas operaciones son realizadas con ayuda del lenguaje Python y bibliotecas como OpenCV.

- Aplicación de la ecualización adaptativa del histograma de contraste limitado (CLAHE) la cual es una técnica basada en la ecualización adaptativa del histograma (AHE), donde el histograma de la imagen resultante se calcula a partir del histograma de una ventana local centrada cada pixel, que se denomina región vecindario. Se ha demostrado que el algoritmo AHE es capaz de realzar zonas de interés en imágenes médicas aunque se produce una sobrealzamiento del ruido. CLAHE divide la imagen en regiones del mismo tamaño sin solapamiento. Tras esto se calcula el histograma de cada una de estas regiones. La amplificación del contraste en cada región de píxeles viene dada por la pendiente de la función de transformación que modifica las intensidades de los píxeles para obtener la imagen de salida. Debido a que CLAHE limita el histograma con un valor predefinido antes de calcular la función de distribución, la pendiente de la función de transformación se verá modificada al fijar el valor límite. 3.7 3.8
- Obtención de la región de interés, aplicando filtros de erosión, detección de bordes, cerradura y creando una mascara binaria para obtener unicamente el área asociada a los pulmones. Este proceso permite a la red neuronal enfocarse en los aspectos relevantes en las imágenes que utilizamos, normalmente tanto en las tomografías como en las radiografías existen elementos que no representan o dan valor al estudio o experimento que estamos realizando, por ejemplo la parte exterior que incluye las costillas y todo lo que se encuentre exterior a ellas resulta de nula utilidad para la detección de características en los pulmones, por este motivo se realiza el proceso en búsqueda de lograr que la red detecte unicamente los elementos que representen y diferencien las clases a las que pertenece cada imagen. 3.9
- Un elemento importante es la generación sintética de imágenes, por lo general este método se utiliza cuando el conjunto de datos es muy reducido, en la mayoría de los proyectos que se revisaron en la bibliografía utilizan este sobre su conjunto de datos debido a las dimensiones de su conjunto de datos, ahora bien el uso de imágenes sintéticas con un conjunto de datos reducido no garantiza el correcto entrenamiento y clasificación de la red neuronal debido a que finalmente son las



FIGURA 3.7: Radiografía de Tórax original sin filtro CLAHE



FIGURA 3.8: Radiografía de Tórax con filtro CLAHE aplicado

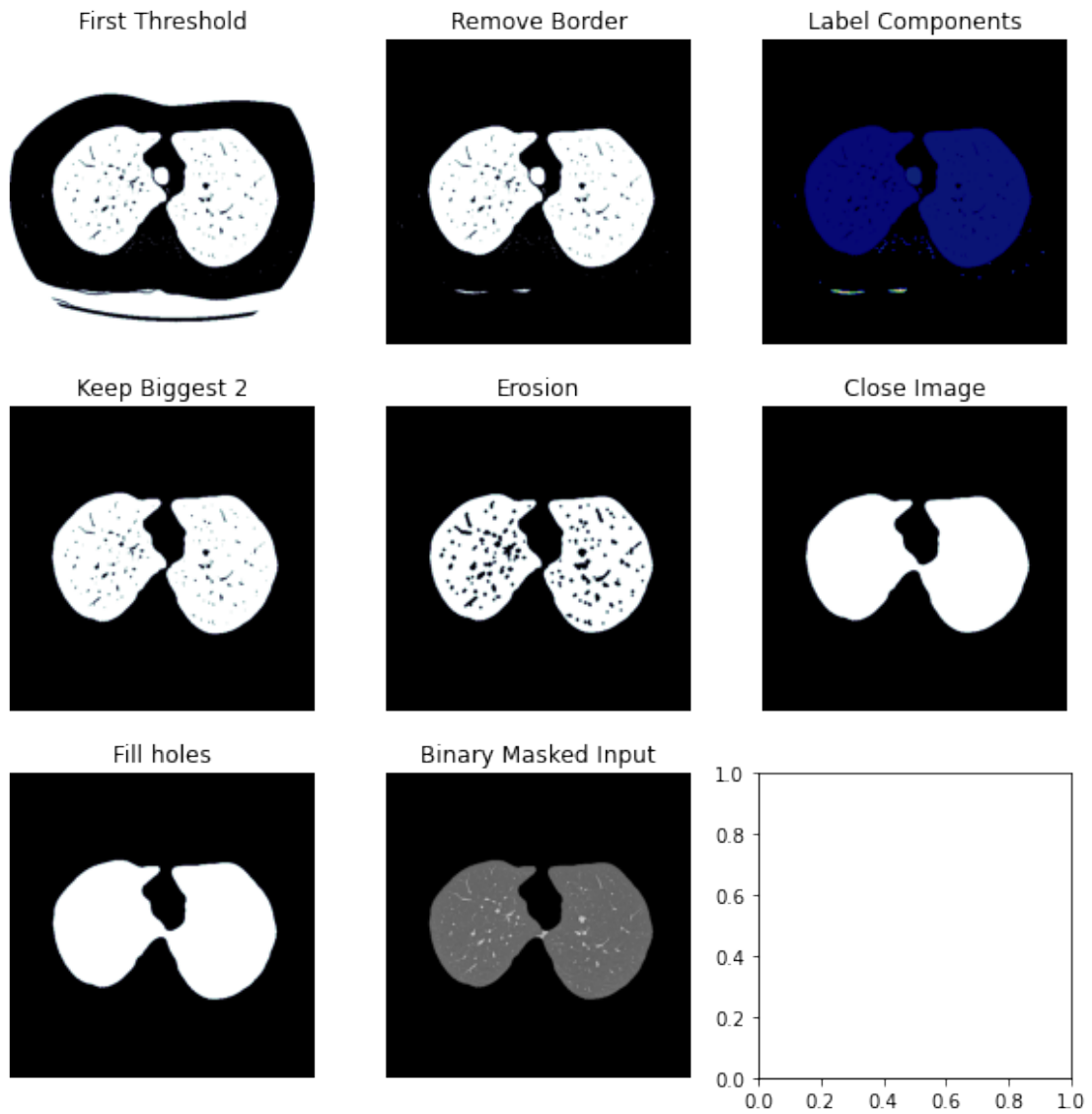


FIGURA 3.9: Preprocesamiento de Región de Interés

mismas imágenes solamente modificadas en ángulos, rotaciones, acercamientos y desplazamientos. Debido a que el conjunto de datos que tenemos es robusto, de dimensiones relativamente grandes, variado respecto a sus fuentes, de alta calidad en su resolución, el uso de la generación sintética de imágenes viene como un elemento adicional al perfeccionamiento de los entrenamientos generando un modelo mejor entrenado, esto significa que generalizara de manera correcta aun cuando las imágenes que ingresen a el no hayan sido adquiridas de forma perfecta o las capacidades de adquisición de los equipos sean diferentes a los que se utilizaron en el proceso de construcción del conjunto de datos que alimento a la red neuronal en su entrenamiento. 3.10

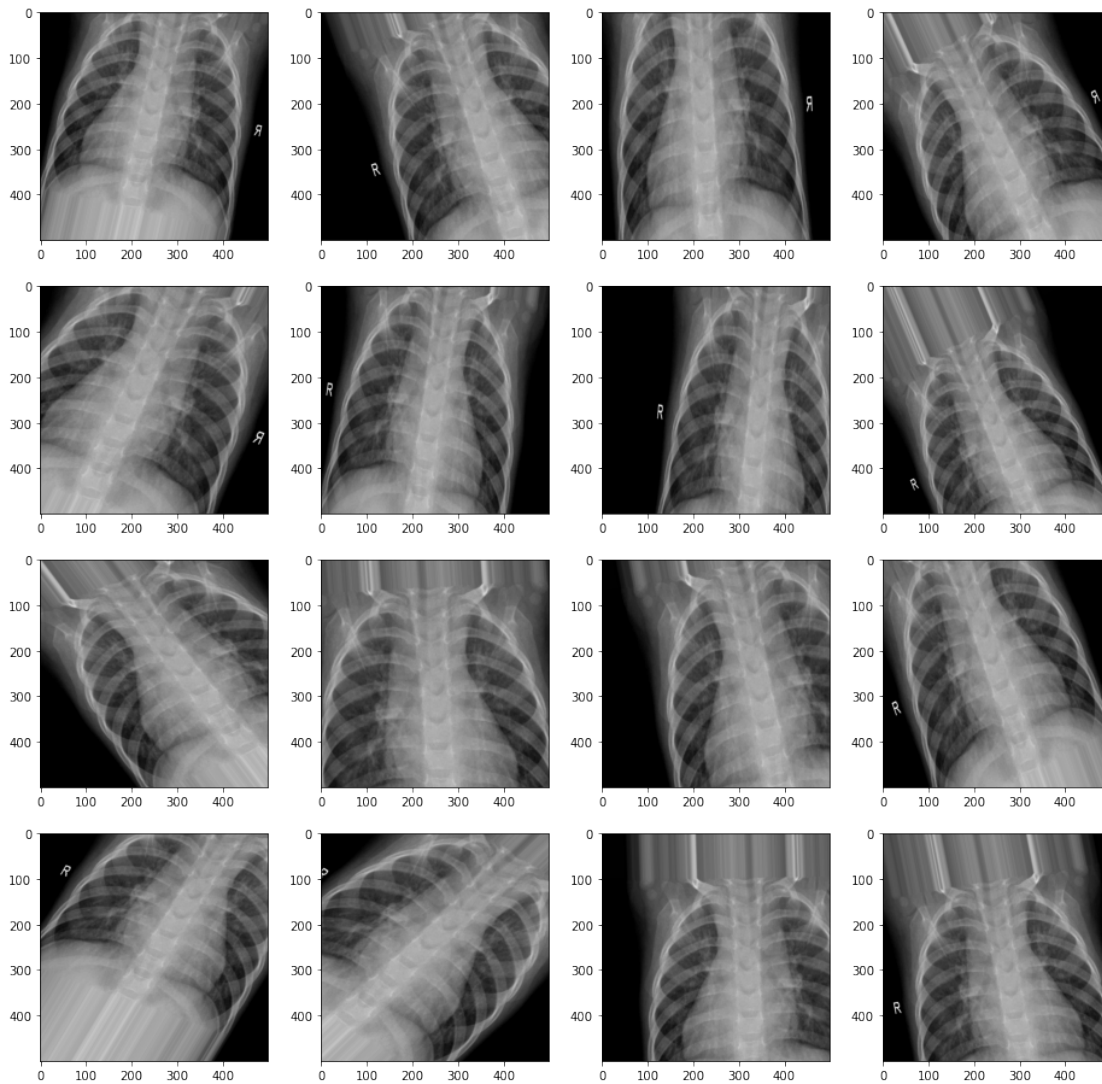


FIGURA 3.10: Generación sintética de imágenes

3.2. Modelo de Red Neuronal Artificial Profunda

La mayoría de los estudios sobre el área de clasificación de imágenes medicas utilizan y confirman que la herramienta que ofrece mejores resultados son las redes neuronales convolucionales (CNN) (Yadav y Jadhav, 2019), estas redes permiten su implementación en diversos lenguajes, uno de ellos el mas utilizado en la actualidad es Python, esto debido a la facilidad de uso, interoperabilidad entre diversos sistemas operativos, arquitecturas de Hardware y dispositivos. Esta fue la razón principal por lo que se decide utilizar este lenguaje.

En esta sección se describe el modelo implementado y las características que se utilizaron para seleccionar las diversas capas de la red neuronal.

3.2.1. Introducción

El uso de métodos tradicionales de aprendizaje maquina como Maquinas de Vector de Soporte (SVM) se han utilizado para la clasificación de imágenes médicas, sin embargo estos tienen ciertas desventajas: el desempeño dista de ser practico y el desarrollo de estas ha sido muy lento en épocas recientes. También la extracción y selección de características consume mucho tiempo y varia de acuerdo al objeto. Las redes neuronales en especial las convolucionales han evolucionado y alcanzado un desempeño significativo desde 2012 (Rawat y Wang, 2017). Algunas investigaciones inclusive han alcanzado grados de desempeños comparables a los humanos.

Los modelos basados en CNN tienen diversas características que permiten incrementar el desempeño, uno de ellos es la generación sintética de imágenes,(Perez y Wang, 2017) Wang y Perez realizaron una investigación sobre el impacto de esta estrategia en la clasificación de imágenes obteniendo mejores resultados que con GAN y otros métodos.

3.2.2. Implementación de modelos de red existentes

Una parte importante del proyecto es analizar los modelos de red existentes para poder realizar una comparativa entre sus niveles de desempeño contra los que se alcancen en este trabajo. Después de realizar un análisis de diversos proyectos que se publicaron al momento de la elaboración de esta investigación. Dentro de las redes mayormente utilizadas encontramos los modelos de red neuronal VGG, RESNET, InceptionV3 y Xception. Estos modelos ya han sido evaluados y probados para el análisis de imágenes. A continuación realizamos una breve descripción de cada uno de ellos.

VGG

VGG (Visual Geometry Group) pertenece al Departamento de Ciencia e Ingeniería de la Universidad de Oxford, y ha lanzado una serie de modelos de redes convolucionales que comienzan con VGG, que pueden aplicarse al reconocimiento de rostros y la clasificación de imágenes, desde VGG16 a VGG19. El propósito original de la investigación de VGG sobre la profundidad de las redes convolucionales es comprender cómo la profundidad de las redes convolucionales afecta la precisión y exactitud de la clasificación y el reconocimiento de imágenes a gran escala. -Deep-16 CNN),Para profundizar el número de capas de red y evitar demasiados parámetros, VGG utiliza un

pequeño núcleo de convolución 3x3 en todas las capas, y el tamaño del paso de la capa de convolución se establece en 1.

La entrada de VGG se establece en una imagen RGB de tamaño 224x224. El valor RGB promedio se calcula para todas las imágenes en la imagen del conjunto de entrenamiento, y luego la imagen se ingresa como una entrada a la red de convolución VGG. Se utiliza un filtro 3x3 o 1x1, y el paso de convolución es fijo. Hay 3 capas VGG completamente conectadas, que pueden variar de VGG11 a VGG19 de acuerdo con el número total de capas convolucionales + capas completamente conectadas. El VGG11 mínimo tiene 8 capas convolucionales y 3 capas completamente conectadas. El VGG19 máximo tiene 16 capas convolucionales. +3 capas totalmente conectadas. Además, la red VGG no es seguida por una capa de agrupación detrás de cada capa convolucional, o un total de 5 capas de agrupación distribuidas bajo diferentes capas convolucionales. Diagrama de estructura:

RESNET

Kaiming He, Xiangyu Zhang, Shaoqin Ren, Jian Sun del equipo de investigación de Microsoft presentaron un marco de aprendizaje residual (ResNets) para ayudar a facilitar el entrenamiento de las redes que son sustancialmente más profundas que antes al eliminar el problema de degradación. Han demostrado con evidencia que los ResNets son más fáciles de optimizar y pueden tener una alta precisión a profundidades considerables. En las redes residuales, en lugar de esperar que las capas se ajusten al mapeo deseado, dejamos que estas capas se ajusten a un mapeo residual.

Este método de omitir los datos de una capa a otra se denomina conexiones de acceso directo o conexiones de omisión. Este enfoque permite que los datos fluyan fácilmente entre las capas sin obstaculizar la capacidad de aprendizaje del modelo de aprendizaje profundo. La ventaja de agregar este tipo de conexión de omisión es que si alguna capa perjudica el rendimiento del modelo, se omitirá.

La intuición detrás de la conexión de salto es que es más fácil para la red aprender a convertir el valor de $f(x)$ a cero para que se comporte como una función de identidad en lugar de aprender a comportarse como una función de identidad por sí sola intentando para encontrar el conjunto correcto de valores que le darían el resultado.

InceptionV3

Inception se enfoca en la extensión. Entonces, los desarrolladores de Inception estaban interesados en la eficiencia computacional de entrenar redes más grandes.

El trabajo original se refería a un componente conocido como el nombre de Inception model. El corazón de este modelo contenía dos elementos innovadores.

La primera intuición se refiere a las operaciones en las capas. En una CNN tradicional, cada capa extrae información de la capa anterior para transformar los datos de origen en un tipo diferente de información. La salida de una convolución de 5x5 extrae características diferentes de las de un 3x3 o un máximo de agrupación, y así sucesivamente.

La mayor densidad de información de la arquitectura de este modelo tiene un problema relevante: el aumento drástico en la capacidad computacional requerida. No solo los filtros convolucionales más grandes (por ejemplo, 5x5) son exigentes y costosos de

calcular, sino que la superposición de diferentes filtros aumenta el número de mapas de características para cada capa. Este es un verdadero impedimento para nuestro modelo.

Digamos que hay M mapas de entrada. Para cada filtro adicional, debemos hacer que se involucre a lo largo de todos los mapas de entrada M ; de la misma manera, si tenemos más N filtros, deberíamos hacer que se convolucionen a lo largo de todos los mapas $N * M$. En otras palabras, cualquier aumento uniforme en el número de filtros da como resultado un aumento al cuadrado en el cálculo. Nuestro módulo Inception ha triplicado o cuadruplicado la cantidad de filtros. Desde un punto de vista computacional, es un problema real.

Esto lleva a la segunda intuición: use circunvoluciones 1×1 para una reducción en la dimensionalidad. Para resolver la dificultad computacional descrita anteriormente, los desarrolladores de Inception usaron convoluciones 1×1 para filtrar la profundidad de las salidas. Estas circunvoluciones tienen en cuenta un valor a la vez, pero a través de múltiples canales, y también pueden extraer información espacial y comprimirla en una dimensión más pequeña. Por ejemplo, con $2 \times 1 \times 1$ filtros, una entrada de tamaño $64 \times 64 \times 100$ (con 100 mapas de características) se puede comprimir a $64 \times 64 \times 20$. Al reducir el número de mapas de entrada, los desarrolladores pudieron superponer las transformaciones de diferentes capas en paralelo, lo que generó redes que eran profundas (con numerosas capas) y amplias (con operaciones paralelas).

la primera versión de Inception, llamada GoogLeNet, es la anteriormente mencionada que tiene 22 capas y que ganaron la competencia ILSVRC 2014. Inception v2 y v3, desarrolladas un año después, son mejores que la versión anterior por varios factores: el más relevante es el refactoring de circunvoluciones grandes en otras más pequeñas y consecutivas, que son más fáciles de aprender. Por ejemplo, en el Inception v3, la convolución 5×5 fue reemplazada por dos convulsiones consecutivas de 3×3 .

El inicio se convirtió rápidamente en una referencia para la arquitectura de modelos posteriores. Su última versión, v4, también agrega conexiones residuales entre cada módulo, creando un híbrido entre Inception y ResNet. No obstante, el factor más importante es que este modelo demuestra la fortaleza de las arquitecturas red-red bien concebidas, que alcanzan un nuevo nivel en la legitimación del poder de las redes neuronales.

Vale la pena señalar que la investigación sobre las siguientes versiones v2 y v3 se publicó el día inmediatamente posterior a la fecha de renuncia en ResNet.

Xception

Xception (eXtreme Inception), con diferencia de las arquitecturas anteriores, ha cambiado nuestra comprensión de las redes neuronales, particularmente las redes convolucionales.

La hipótesis fue: las correlaciones espaciales y las que existen entre los canales son lo suficientemente libres para hacer que sea preferible no mapearlas juntas.

Esto significa que en una CNN tradicional, los estados convolucionales identificaron correlaciones entre el espacio y la profundidad. En la siguiente imagen podemos revisar la operación.

El filtro considera simultáneamente una dimensión espacial (cada cuadrado de 2×2 colores) y un cross-channel o dimensión de profundidad (la pila de cuatro cuadrados).

En la capa de entrada de una imagen, esto es el equivalente de un filtro convolucional que examina un grupo de píxeles de 2×2 que atraviesa los tres canales RGB.

Inception comienza a separar estos dos elementos: gracias a las convoluciones 1×1 proyecta la entrada original en múltiples espacios de entrada de menor tamaño, y de cada uno de estos utiliza diferentes filtros para transformar esos pequeños bloques de datos tridimensionales. Xception va aún más allá: en lugar de dividir los datos de entrada en múltiples grupos comprimidos, mapea separadamente para realizar una convolución 1×1 en profundidad para capturar correlaciones cross-channel.

Esta operación se conoce como Depthwise Separable Convolution que consiste en una convolución espacial (Depthwise Convolution) realizada independientemente para cada canal, seguida de una convección 1×1 (Pointwise Convolution) entre los canales. Por lo tanto, es una búsqueda de correlaciones primero en un espacio bidimensional y luego en un espacio unidimensional. Esta asignación $2D + 1D$ es más fácil de aprender que una completamente $3D$.

Xception supera el rendimiento de Inception v3 en los conjuntos de datos de ImageNet y lo hace igualmente bien con los conjuntos de datos para la clasificación de imágenes de 17,000 clases. Además, tiene la misma cantidad de parámetros del modelo Inception. Aunque Xception es muy reciente, su arquitectura ya se ha convertido en parte de las aplicaciones de Google para visión móvil, a través de MobileNet.

3.2.3. Diseño

El diseño general del modelo que se utiliza para la clasificación de las imágenes que se realiza en este trabajo incluye el procesamiento de las estas para alimentar el modelo de red que se propone y describe a continuación. 3.11

Inicialmente se obtienen los conjuntos de datos, estos deben de estar en carpetas train, val y test, cada una de ellas nos servira para realizar el entrenamiento, validación y prueba respectivamente. posteriormente se realiza el preprocesamiento de las imágenes y se remplazan periódicamente en cada una de las carpetas, una vez listas se alimenta la red neuronal y se inicia el entrenamiento.

El diseño de la red fue logrado despues de realizar numerosas pruebas con varios modelos de redes existentes y tratando de obtener las mejores características de cada uno de ellos, se utilizaron kernels de 3×3 para la realización de la convolución, se noto que con kernels mas grandes no se lograba capturar correctamente las características que identifican a la enfermedad, finalmente el modelo se muestra en la figura 3.12.

3.2.4. Implementación

Hardware

Para la implementación y entrenamiento de la red propuesta se utilizo un equipo de reciente generación de procesamiento de alto desempeño llamado IBM AC922 V100 el cual permite la aceleración de manera substancial en la ejecución de los experimentos. Las características específicas del equipo son descritas a continuación:

- 2 procesadores Power9 con 40 nucleos
- 1 TB de memoria RAM DDR4 2666 MHz

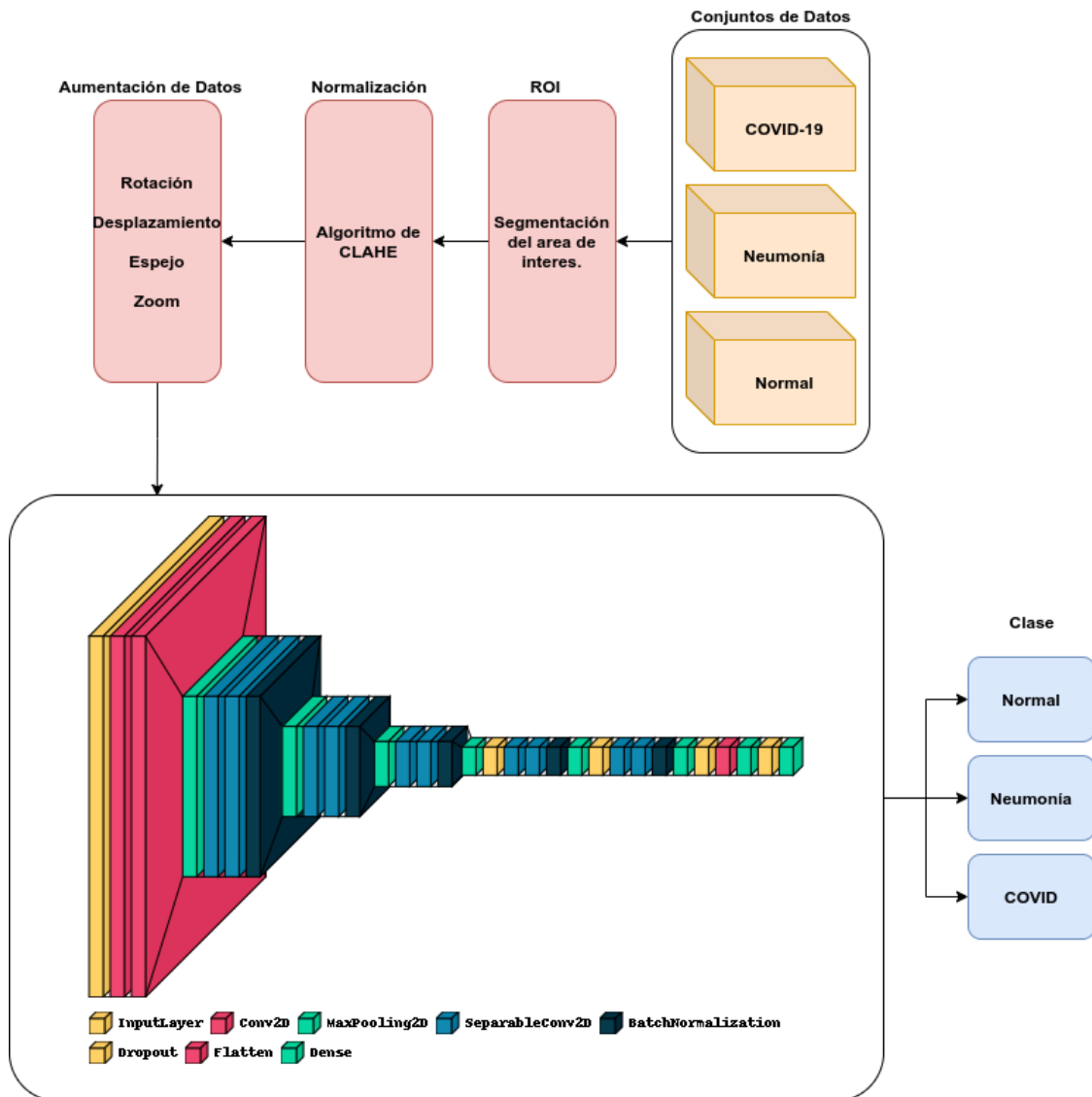


FIGURA 3.11: Modelo completo de procesamiento y clasificación

FIGURA 3.12: Modelo de red CNN propuesto

```

Model: "model"

```

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 256, 256, 3)]	0
conv2d_4 (Conv2D)	(None, 256, 256, 16)	448
conv2d_5 (Conv2D)	(None, 256, 256, 16)	2320
max_pooling2d_2 (MaxPooling2D)	(None, 128, 128, 16)	0
separable_conv2d (SeparableConv2D)	(None, 128, 128, 32)	688
separable_conv2d_1 (SeparableConv2D)	(None, 128, 128, 32)	1344
batch_normalization (Batch Normalization)	(None, 128, 128, 32)	128
max_pooling2d_3 (MaxPooling2D)	(None, 64, 64, 32)	0
separable_conv2d_2 (SeparableConv2D)	(None, 64, 64, 64)	2400
separable_conv2d_3 (SeparableConv2D)	(None, 64, 64, 64)	4736
batch_normalization_1 (Batch Normalization)	(None, 64, 64, 64)	256
max_pooling2d_4 (MaxPooling2D)	(None, 32, 32, 64)	0
separable_conv2d_4 (SeparableConv2D)	(None, 32, 32, 128)	8896
separable_conv2d_5 (SeparableConv2D)	(None, 32, 32, 128)	17664
batch_normalization_2 (Batch Normalization)	(None, 32, 32, 128)	512
max_pooling2d_5 (MaxPooling2D)	(None, 16, 16, 128)	0
dropout (Dropout)	(None, 16, 16, 128)	0
separable_conv2d_6 (SeparableConv2D)	(None, 16, 16, 256)	34176
separable_conv2d_7 (SeparableConv2D)	(None, 16, 16, 256)	68096
batch_normalization_3 (Batch Normalization)	(None, 16, 16, 256)	1024
max_pooling2d_6 (MaxPooling2D)	(None, 8, 8, 256)	0
dropout_1 (Dropout)	(None, 8, 8, 256)	0
separable_conv2d_8 (SeparableConv2D)	(None, 8, 8, 512)	133888
separable_conv2d_9 (SeparableConv2D)	(None, 8, 8, 512)	267264
batch_normalization_4 (Batch Normalization)	(None, 8, 8, 512)	2048
max_pooling2d_7 (MaxPooling2D)	(None, 4, 4, 512)	0
dropout_2 (Dropout)	(None, 4, 4, 512)	0
flatten (Flatten)	(None, 8192)	0
dense (Dense)	(None, 512)	4194816
dropout_3 (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 128)	65664
dropout_4 (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 3)	387

```

Total params: 4,806,755
Trainable params: 4,804,771
Non-trainable params: 1,984

```

- 4 tarjetas NVIDIA Tesla V100 con 16 GB HBM2 de memoria con 900 Gbps de ancho de banda.
- capacidad de 125 TeraFlops para procesamiento con aplicaciones que utilicen aprendizaje profundo.

El uso de este equipo permitió realizar diversos ensayos que hubieran sido prácticamente imposibles de realizar en una computadora de escritorio o similar. Dentro de las ventajas que ofrece esta arquitectura es la integración de un software de desarrollo especializado para este tipo de equipos que provee el fabricante el cual se llama powerai, dentro de este se incluye todo un conjunto de bibliotecas para el desarrollo y aceleración de múltiples aplicaciones entre ellas el aprendizaje profundo. La característica principal que tiene este equipo es la capacidad de comunicar las tarjetas NVidia por medio de un enlace NVLink el cual incrementa el ancho de banda de comunicación hasta los 300 Mbps, además cada tarjeta cuenta con 5120 núcleos CUDA y 520 núcleos Tensor, en total más de 20,000 núcleos CUDA son utilizables y hasta 2040 núcleos Tensor. Este tipo de tecnología se encuentra a la vanguardia y permite la aceleración de la ejecución de aplicaciones de aprendizaje profundo de manera significativa,

Software

Una parte fundamental del proyecto es la correcta selección de los elementos de software que se utilizaran para la implementación, como sistema operativo se utiliza Linux, esto debido a su estabilidad y compatibilidad con software científico. El equipo IBM tiene instalado la versión RedHat 7 lo que permite un soporte aun mayor en caso de incidencias y depuración respecto al sistema operativo.

Se utilizó el entorno de desarrollo de Tensorflow con el lenguaje de programación Python, son dos de las tecnologías mayormente utilizadas para ciencia de datos, aprendizaje máquina y aprendizaje profundo, proveen de una cantidad bastante significativa de

3.2.5. Entrenamiento

Para realizar los entrenamientos de la red neuronal se configuró el sistema para que cubriera estas características:

- De los conjuntos de entrenamiento tanto de rayos X como de CT se utilizó del total del conjunto de datos el 60% para entrenamiento, 20% para validación y 20% para prueba. Antes de seleccionar los conjuntos se realiza una distribución aleatoria del orden de las imágenes para que no influya de manera errónea el nombre que puedan tener los archivos.
- Se establece un rango de aprendizaje en 0.0001, es un valor que permite un crecimiento controlado del aprendizaje sin que se tienda al sobreentrenamiento.
- Se define un máximo de 100 épocas por entrenamiento, esto debido al tiempo que tarda en ejecutarse cada uno de los entrenamientos.

- La resolución que se utiliza es de 512×512 , normalmente la mayoría de los sistemas definen una resolución de hasta 256×256 , pero debido a que se cuenta con un equipo de características superiores es deseable explotar en su totalidad sus capacidades. La mejora en la resolución permitirá que la detección de las características sea mucho mejor.
- Adam es un algoritmo de optimización que se puede utilizar en lugar del procedimiento de descenso de gradiente estocástico clásico para actualizar los pesos de red de forma iterativa en función de los datos de entrenamiento. El descenso de gradiente estocástico mantiene una única tasa de aprendizaje (denominada alfa) para todas las actualizaciones de peso y la tasa de aprendizaje no cambia durante el entrenamiento. El método calcula las tasas de aprendizaje adaptativo individuales para diferentes parámetros a partir de estimaciones del primer y segundo momento de los gradientes.

En lugar de adaptar las tasas de aprendizaje de los parámetros en función del primer momento promedio (la media) como en RMSProp, Adam también hace uso del promedio de los segundos momentos de los gradientes (la varianza no centrada).

Específicamente, el algoritmo calcula un promedio móvil exponencial del gradiente y el gradiente al cuadrado, y los parámetros beta1 y beta2 controlan las tasas de caída de estos promedios móviles.

El valor inicial de las medias móviles y los valores beta1 y beta2 cercanos a 1.0 (recomendado) dan como resultado un sesgo de las estimaciones de momento hacia cero. Este sesgo se supera calculando primero las estimaciones sesgadas antes de calcular luego las estimaciones corregidas por sesgo.

El algoritmo de Adam obtiene las ventajas de los algoritmos AdaGrad y RMSProp. Adam no solo calcula la tasa de aprendizaje de parámetros adaptativos en función del valor medio del primer momento como el algoritmo RMSProp, sino que también hace un uso completo del valor medio del segundo momento del gradiente (es decir, la varianza no centrada). Específicamente, el algoritmo calcula el promedio móvil exponencial del gradiente, y los hiperparámetros beta1 y beta2 controlan la tasa de disminución de estos promedios móviles. El valor inicial de la media móvil y los valores de beta1 y beta2 están cerca de 1 (valor recomendado), por lo que la desviación de la estimación de momento está cerca de 0. La desviación se mejora calculando primero la estimación con desviación y luego calculando la estimación después de la corrección de desviación.

- Se utiliza para todos los casos un algoritmo de validación cruzada en el cual el conjunto de entrenamiento fue dividido en 60% para entrenamiento, 20% para validación y 20% para prueba, el segmento de validación es modificado en cada época de entrenamiento realizando una variación aleatoria de los datos, entonces en cada una de las épocas el conjunto de validación y entrenamiento es diferente de esta forma el sistema logra generalizar de forma mas eficiente.

Capítulo 4

Capitulo 4 - Análisis de Resultados

4.1. Introducción

Es importante contemplar que cualquier sistema que involucre un análisis y clasificación de imágenes de pacientes con cualquier enfermedad siempre tendrá que ser validado en algún momento por personal capacitado para esta tarea. En ese sentido la principal característica que buscamos en el sistema

Estas métricas también tienen su correspondiente nombre en español, pero es importante que sepas su nombre en inglés porque muchas librerías (scikit-learn), las tienen ya implementadas.

La medida más deseada para este sistema es lograr que se clasifiquen la mayor cantidad de personas con una posible afectación por COVID-19 en los pulmones, existirán casos en los que el sistema genere una "falsa alarma" por ejemplo una persona sana o con neumonía de otro tipo y sean clasificados como COVID-19, pero no sería en ningún momento aceptable clasificar a una persona con COVID-19 como normal. En consecuencia la métrica que resulta más interesante para el proyecto es la de recall o sensibilidad, más adelante se detalla como es que se obtiene su valor.

4.2. Métricas de Evaluación

Estas métricas también tienen su correspondiente nombre en español, pero es importante que utilice su nombre en inglés porque muchas librerías (scikit-learn), las tienen ya implementadas. En la siguiente tabla se hace referencia a su correspondencia con el idioma inglés.

En el campo de la inteligencia artificial y el aprendizaje automático una matriz de confusión es una herramienta que permite visualizar el desempeño de un algoritmo de aprendizaje supervisado. Cada columna de la matriz representa el número de predicciones de cada clase, mientras que cada fila representa a las instancias en la clase real, o sea en términos prácticos nos permite ver qué tipos de aciertos y errores está teniendo nuestro modelo a la hora de pasar por el proceso de aprendizaje con los datos.

Pensemos en la clasificación de pacientes con COVID-19 en positivos y negativos:

En este ejemplo tenemos un grupo de 100 pacientes de los cuales hay 45 que sí tienen el virus y 55 que no lo tienen.

Nuestro algoritmo supervisado de clasificación entre otras cosas, nos va a ayudar a determinar cuál es el % de acierto de nuestras pruebas.

Realizamos las pruebas respectivas a los pacientes y de esta forma, podríamos comparar los resultados del modelo real contra el modelo predictivo:

Inglés	Español
Precision	Precisión
Recall	Sensibilidad
F1-score	Valor-F
Accuracy	Exactitud
Confusion Matrix	Matriz de Confusión
True Positive	Positivos Verdaderos
True Negative	Negativos Verdaderos
False Positive	Positivos Falsos
False Negative	Negativos Falsos

CUADRO 4.1: Referencia de traducción de términos relacionados con las métricas utilizadas en este proyecto.

Las 4 opciones siguientes son las que conforman lo que se conoce como la matriz de confusión:

en este caso al ser sólo dos posibilidades: positivo o negativo (o clase 1 y clase 2]), hablamos de una matriz binaria. Entonces surgen cuatro opciones:

Persona que tiene COVID-19 y el modelo lo clasificó como COVID-19 (+) . Esto sería un verdadero positivo o VP . Persona que no tiene COVID-19 y el modelo lo clasificó como COVID-19 (-) . Este sería un verdadero negativo o sea un VN. Persona que tiene COVID-19 y el modelo lo clasificó como COVID-19 (-) . Éste sería un falso negativo o FN. Persona que no tiene COVID-19 y el modelo lo clasificó como COVID-19 (+) . Este es un falso positivo o FP. Ahora de forma más clara, podemos identificar en nuestra matriz donde se ubican los errores.

4.2.1. Exactitud

la Exactitud (en inglés, “Accuracy”) se refiere a lo cerca que está el resultado de una medición del valor verdadero. En términos estadísticos, la exactitud está relacionada con el sesgo de una estimación. Se representa como la proporción de resultados verdaderos (tanto verdaderos positivos (VP) como verdaderos negativos (VN)) dividido entre el número total de casos examinados (verdaderos positivos, falsos positivos, verdaderos negativos, falsos negativos)

En forma práctica, la Exactitud es la cantidad de predicciones positivas que fueron correctas. La fórmula para realizar el cálculo es la siguiente:

$$\frac{VP + VN}{VP + FP + FN + VN}$$

4.2.2. Precisión

La Precisión (en inglés “Precision”) Se refiere a la dispersión del conjunto de valores obtenidos a partir de mediciones repetidas de una magnitud. Cuanto menor es la dispersión mayor la precisión. Se representa por la proporción de verdaderos positivos dividido entre todos los resultados positivos (tanto verdaderos positivos, como falsos positivos).

En forma práctica es el porcentaje de casos positivos detectados. La Forma para calcularla es la siguiente:

$$\frac{VP}{VP + FP}$$

4.2.3. Sensibilidad y la Especificidad

La sensibilidad y la especificidad son dos valores que nos indican la capacidad de nuestro estimador para discriminar los casos positivos, de los negativos. La sensibilidad se representa como la fracción de verdaderos positivos, mientras que la especificidad, es la fracción de verdaderos negativos.

La Sensibilidad ("Recall" o "Sensitivity"), También se conoce como Tasa de Verdaderos Positivos (True Positive Rate) ó TP. Es la proporción de casos positivos que fueron correctamente identificadas por el algoritmo. Se calcula así :

$$\frac{VP}{VP + FN}$$

o lo que sería igual : Verdaderos positivos / Total Enfermos En el área de la salud decimos que la sensibilidad es la capacidad de poder detectar correctamente la enfermedad entre los enfermos.

La Especificidad ("Specificity") También conocida como la Tasa de Verdaderos Negativos, ("true negative rate") o TN. Se trata de los casos negativos que el algoritmo ha clasificado correctamente. Expresa cuan bien puede el modelo detectar esa clase. Se calcula:

$$\frac{VN}{VN + FP}$$

En términos de salud: Verdaderos Negativos / Total Sanos En el área de la salud decimos que la especificidad es la capacidad de poder identificar los casos de pacientes sanos entre todos los sanos

4.2.4. ROC

La curva ROC es una herramienta estadística utilizada en el análisis de la clasificar la capacidad discriminante de una prueba diagnóstica dicotómica. Es decir, una prueba, basada en una variable de decisión, cuyo objetivo es clasificar a los individuos de una población en dos grupos: uno que presente un evento de interés y otro que no. Esta capacidad discriminante está sujeta al valor umbral elegido de entre todos los posibles resultados de la variable de decisión , es decir, la variable por cuyo resultado se clasifica a cada individuo en un grupo u otro. La curva es el gráfico resultante de representar, para cada valor umbral, las medidas de sensibilidad y especificidad de la prueba diagnóstica. Por un lado, la sensibilidad cuantifica la proporción de individuos que presenta el evento de interés y que son clasificados por la prueba como portadores de dicho evento. Por otro lado, la especificidad cuantifica la proporción de individuos que no lo presentan y son clasificados por la prueba como tal.

Las ventajas del uso de la curva ROC son: que proporciona una representación de la sensibilidad y especificidad para cada valor umbral, que es invariante mediante transformaciones monótonas a los datos de la variable de decisión y que permite comparar dos o más clasificadores en función de su capacidad discriminante.

4.3. Resultados del modelo implementado

El modelo que se diseñó e implementó en este proyecto presenta resultados satisfactorios debido a que alcanza una precisión que permite su utilización en el ambiente médico por medio de la determinación de pacientes con una alta posibilidad de neumonía causada por COVID-19, es de relevancia hacer el señalamiento que este desarrollo requiere aun de una optimización, refinamiento y ajustes que exceden los alcances de este, pero es un principio para la utilización de mecanismos que permitan extender su funcionamiento.

El uso de diversos conjuntos de datos así como la generación de un conjunto de datos propio, permite que el modelo de red CNN se entrene con una diversidad de imágenes que enriquece y valida los resultados de los modelos, muchos de los estudios y proyectos iniciales que se realizaron sobre el análisis de imágenes de pacientes con COVID-19 obtenían precisiones desde el 80 por ciento hasta el 95 por ciento pero estos resultados se veían sesgados al momento de introducir imágenes con características un poco diferentes, el motivo principal de este sesgo fue el uso de conjuntos de datos muy reducidos, lo que provocaba un sobre ajuste provocando poca generalización del modelo. Posteriormente conforme avanzó la enfermedad ya se empezó a contar con mayor disponibilidad de datos, aun así son relativamente pocos los proyectos que utilizan una dimensionalidad de datos como la que se utilizó en nuestra investigación.

4.3.1. Matriz de Confusión

La matriz de confusión nos permite observar que al momento de ejecutar el sistema de pruebas de la red neuronal esta logra clasificar de manera bastante exitosa a los casos de COVID-19 por medio de Tomografías con un porcentaje del 99%. Por medio de la matriz de confusión podemos observar como el modelo que se propone tiene resultados bastante satisfactorios sobre todo en la detección y clasificación de los pacientes con afectaciones en los pulmones relacionada con COVID-19. Este resultado es especialmente satisfactorio para el modelo que clasifica tomografías, esto debido a que como se estudio en la bibliografía, la tomografía ofrece una mayor cantidad de detalles que permiten realizar un proceso de diagnóstico tanto humano como por aprendizaje máquina más preciso y eficiente.

4.3.2. Reporte de Clasificación

Este reporte muestra los resultados que se tuvieron durante el proceso de entrenamiento, recordemos que la parte de validación se realiza con el 20% del total de cada uno de los conjuntos de datos de entrenamiento. Este resultado está directamente relacionado con la matriz de confusión debido a que a partir de ella es como se generó la matriz, uno de los valores que tienen mayor relevancia es la existencia de un recall o

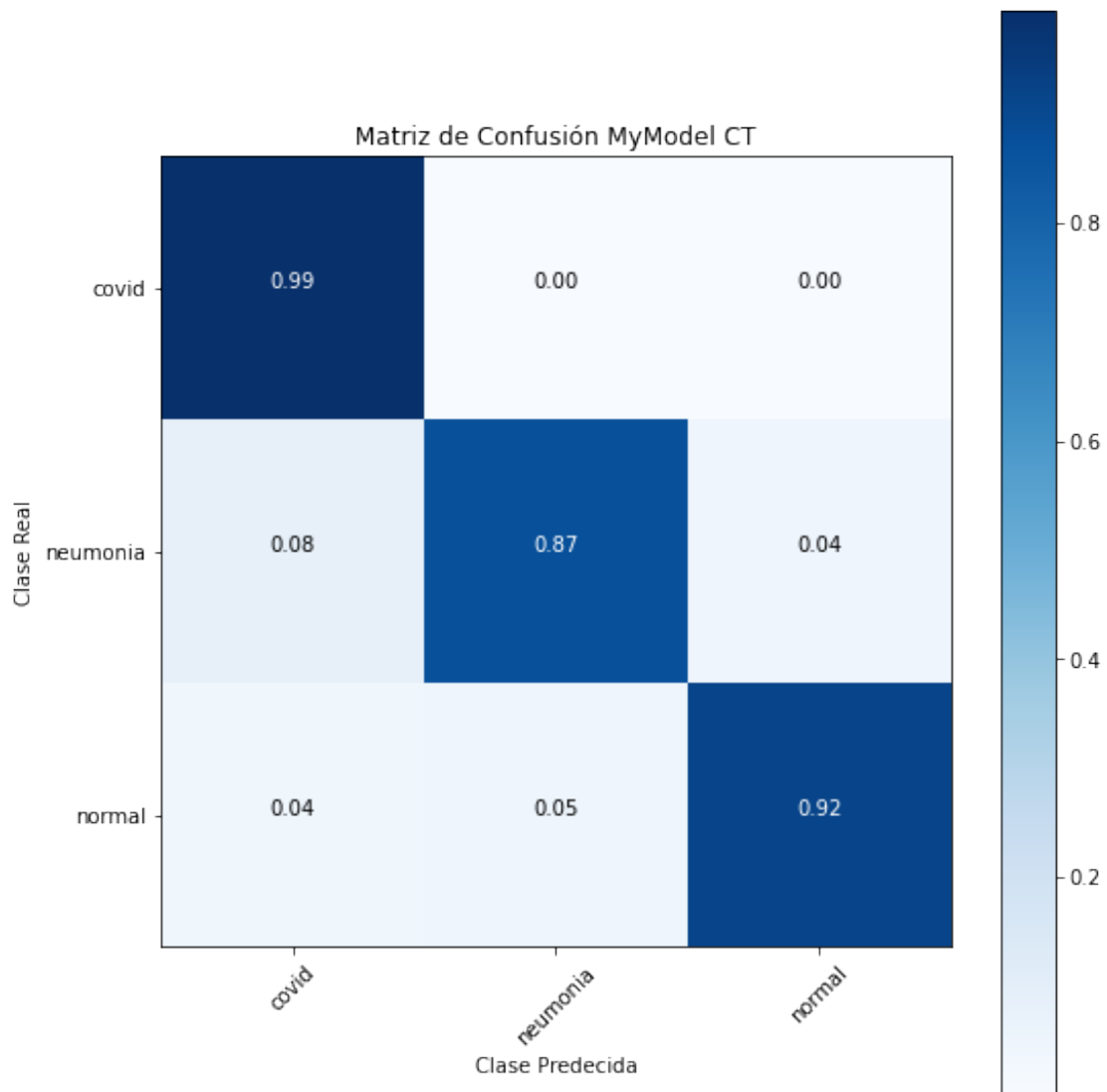


FIGURA 4.1: Matriz de Confusión para Nuestro modelo con CT

sensibilidad muy alto para el caso de COVID-19, esto significa que un porcentaje mínimo de personas con COVID-19 serán clasificadas por el sistema de otra forma, también los porcentajes de precisión de las demás clases resultan alentadores e indican que el proceso de entrenamiento y validación se realizó de forma correcta. ??

Informe de clasificación MyModel CT				
	precision	recall	f1-score	support
covid	0.89	0.99	0.94	6244
neumonía	0.95	0.87	0.91	7400
normal	0.95	0.92	0.93	6842
accuracy			0.93	20486
macro avg	0.93	0.93	0.93	20486
weighted avg	0.93	0.93	0.93	20486

CUADRO 4.2: Reporte de Clasificación para el modelo propio con datos CT

4.3.3. la Curva ROC

Recordando que la curva ROC presenta una visión gráfica del comportamiento que tienen los modelos respecto a la clasificación de un grupo de imágenes contra las clasificaciones erróneas que se realizan, podemos observar que el modelo que se entreno tiene un buen valor de AUC (Área Bajo la Curva) lo cual es un indicador que la clasificación se realizó de manera correcta cuando al sistema se le ingresen imágenes tomográficas. Los valores AUC de neumonía y normal también son satisfactorios y como se verá más adelante sobrepasan de manera significativa a los resultados de los modelos tradicionales. 4.2

4.4. Comparativa con los modelos Inception, VGG, ResNet, Xception

Los resultados de clasificación de los modelos tradicionalmente utilizados para la clasificación de imágenes, dan una comparativa real contra la cual podemos validar nuestro modelo, nos presentan una perspectiva del funcionamiento en la vida real alcanzado por el modelo propuesto. como se ve en ellos todos obtuvieron valores por debajo del implementado en este proyecto, el que tiene un mejor resultado es el modelo de ResNet el cual nos entrega una precisión del 80 % y un recall del 78 % los cuales están muy distantes de los resultados que se obtuvieron con el modelo, aun así nos sirven de parámetro y puede significar que la utilización de este modelo de red con ciertas modificaciones y variaciones tendría resultados mejores, pero esta prueba excede los alcances de este proyecto. 4.3

Las matrices de confusión de los modelos comparativos nos demuestran de forma gráfica como es que nuestro modelo obtiene resultados muy alentadores y es indicativo que el modelo está correctamente implementado y entrenado, en estas gráficas podemos

Informe de clasificación Inception CT				
	precision	recall	f1-score	support
covid	0.87	0.59	0.70	6244
neumonia	0.73	0.93	0.82	7400
normal	0.91	0.90	0.91	6842
accuracy			0.82	20486
macro avg	0.84	0.81	0.81	20486
weighted avg	0.83	0.82	0.81	20486
Informe de clasificación ResNet CT				
	precision	recall	f1-score	support
covid	0.85	0.78	0.81	6244
neumonia	0.81	0.92	0.86	7400
normal	0.91	0.85	0.88	6842
accuracy			0.85	20486
macro avg	0.86	0.85	0.85	20486
weighted avg	0.86	0.85	0.85	20486
Informe de clasificación VGG				
	precision	recall	f1-score	support
covid	0.51	0.39	0.44	6244
neumonia	0.61	0.66	0.63	7400
normal	0.76	0.85	0.80	6842
accuracy			0.64	20486
macro avg	0.63	0.63	0.63	20486
weighted avg	0.63	0.64	0.63	20486
Informe de clasificación Xception CT				
	precision	recall	f1-score	support
covid	0.80	0.67	0.73	6244
neumonia	0.76	0.87	0.81	7400
normal	0.86	0.86	0.86	6842
accuracy			0.81	20486
macro avg	0.81	0.80	0.80	20486
weighted avg	0.81	0.81	0.80	20486

CUADRO 4.3: Reportes de Clasificación de los modelos comparativos.

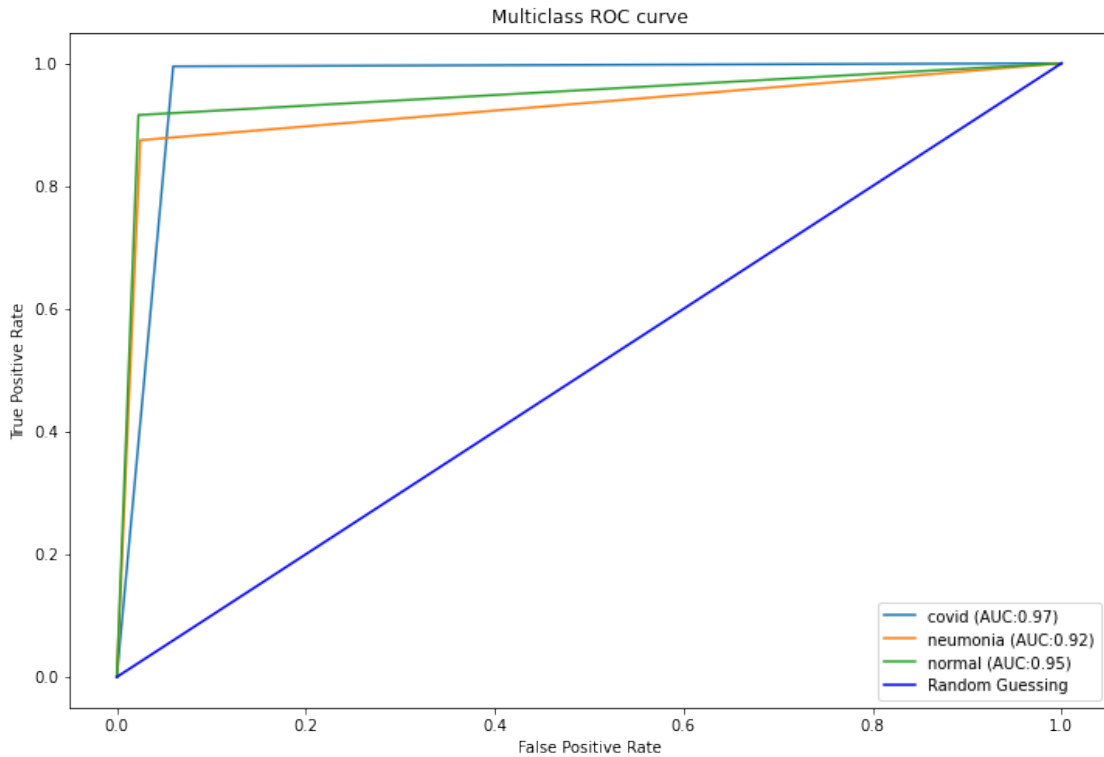
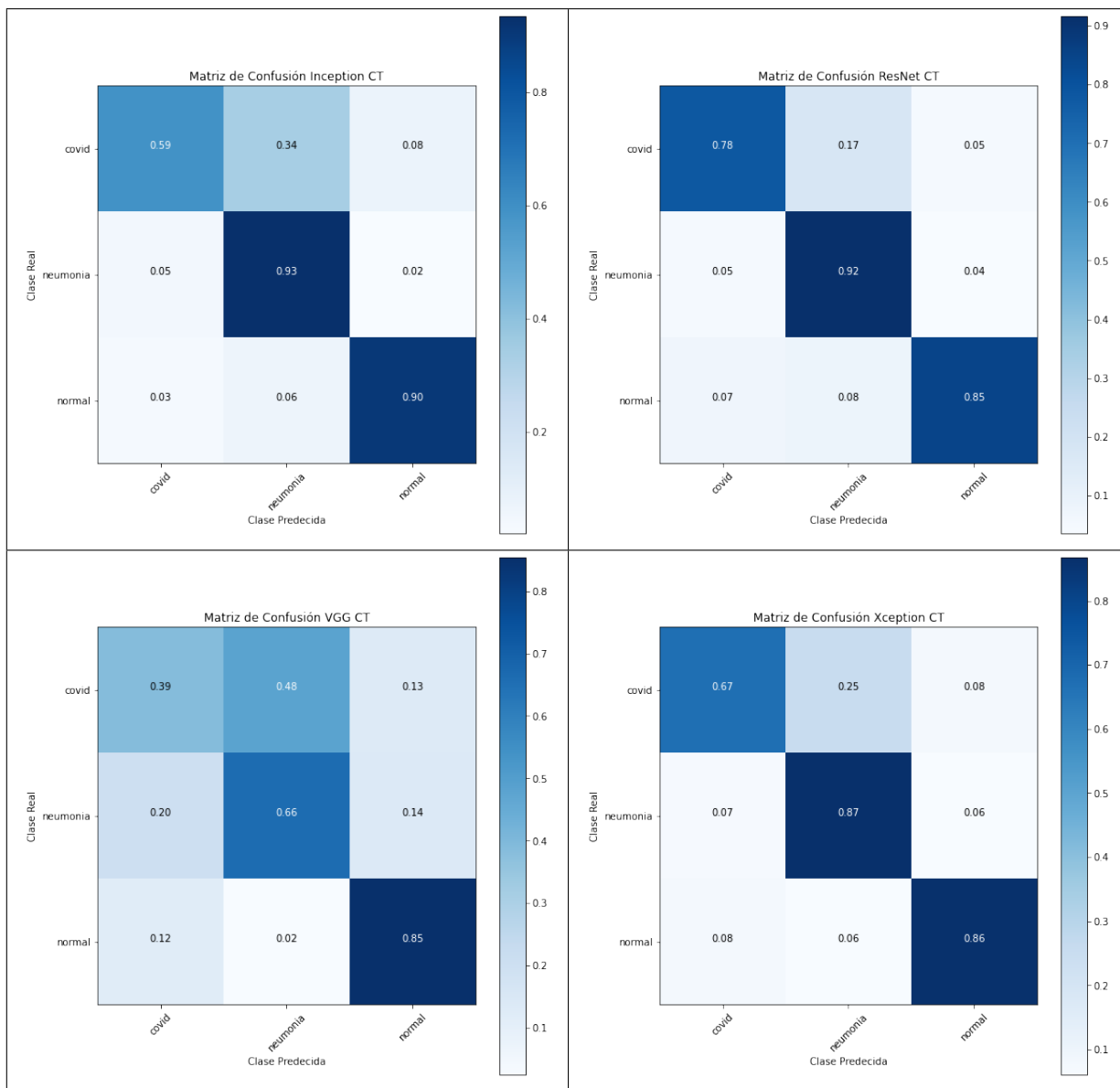


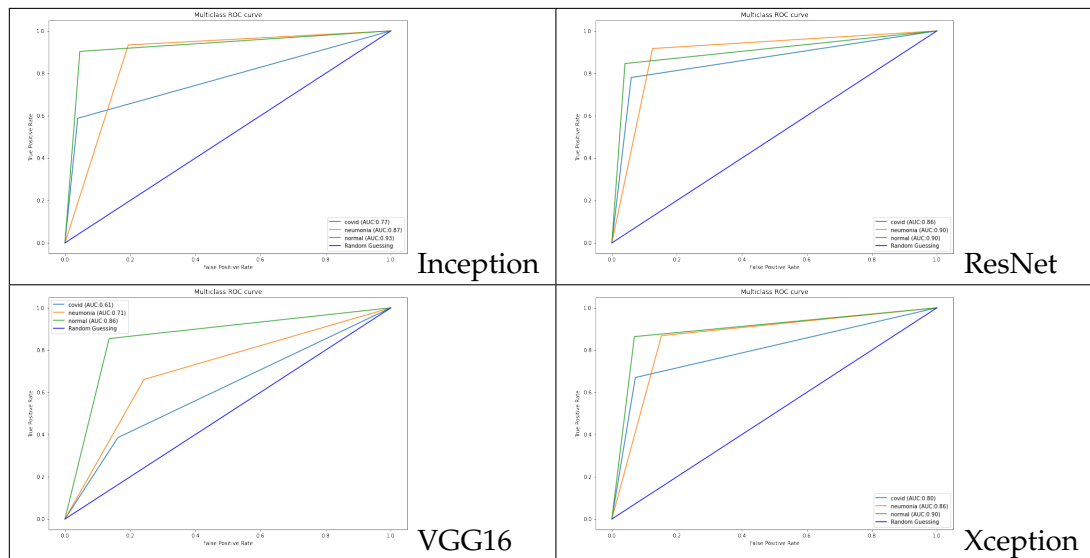
FIGURA 4.2: ROC para el modelo entrenado con Tomografías

observar que todos ellos obtienen buenos resultados el momento de detectar pacientes sanos, el problema radica en la deserción de las clases de neumonía y COVID-19, lo cual es realmente complicado debido a la similitud que existe en las afectaciones que tienen estas dos clases. Digamos que siempre sera mas sencillo diferenciar dos imágenes cuando los pacientes son sanos y enfermos. La diferencia mas significativa se encuentra en la parte de COVID-19 donde encontramos que ninguno logra una clasificación lo suficientemente satisfactoria para poder ser utilizado en algún área médica debido a los valores que se tienen de error de clasificación por lo cual muchos pacientes con neumonias causadas por COVID-19 serian clasificados como neumonía de otro tipo o inclusive sanos lo que resultaría en una falla garrafal del sistema. 4.4

Los curvas de ROC de los modelos que se utilizaron para comparar contra el modelo que proponemos también muestran un desempeño pobre respecto a las clasificación de las imágenes, en cada una de ellas podemos observar un desempeño bastante inferior y de igual forma se hace notable la clasificación errónea que hacen en los casos de COVID-19 y neumonía, en casi todas ellas, a excepción de VGG16, la AUC de de normal tiene valores aceptables pero las de neumonía y COVID-19 resultan muy bajas provocando así que en situaciones de la vida real la clasificación por medio de estas redes resulte de poca o nula ayuda al personal médico.



CUADRO 4.4: Matrices de Confusión de los modelos de Red comúnmente utilizados en imágenes, utilizando tomografías .



CUADRO 4.5: Curvas ROC de modelos de Red comúnmente utilizados en imágenes, utilizando tomografías

Capítulo 5

Capitulo 5 - Conclusiones

5.1. Conclusiones

La enfermedad COVID-19 causada por el virus SARSCoV2 persiste hasta el momento de la terminación de este proyecto, múltiples variantes han surgido provocando diversas mutaciones del virus las cuales principalmente se diferencian, hasta el momento, por la contagiabilidad y persistencia en el medio ambiente pero la sintomatología y afectaciones a los pulmones permanecen con la misma consistencia. Es evidente que tendremos que acostumbrarnos a vivir con ella, sin embargo la producción de herramientas que permitan un diagnostico temprano sigue siendo importante para la conservación de la salud de la población en general.

Por medio de aplicaciones y metodologías como la que se muestra en este trabajo es posible ayudar al personal médico al diagnostico temprano de la enfermedad, siempre con ayuda de un experto que evalúe y valide los resultados que se obtengan del proceso. Existen en desarrollo múltiples modelos, aplicaciones y técnicas para la mejora, optimización y perfeccionamiento de la detección, clasificación y seguimiento de pacientes con afectaciones pulmonares ocasionados por la enfermedad.

El uso de aprendizaje profundo en específico de las redes CNN es un área relativamente reciente y aun se encuentra en desarrollo pero ha demostrado con resultados muy alentadores su eficacia en los procesos que involucran imágenes medicas, en el caso específico de CT y RX de tórax logran resolver de manera satisfactoria problemáticas en el ámbito de la salud que anteriormente hubiera sido de una complejidad mayor resolver.

Es notorio que la mayoría de los diagnósticos por imágenes de tórax involucran datos obtenidos por medio de tomografías computarizadas, esto tiene sentido debido a que desde un inicio fue mayor la probabilidad y facilidad de detección de las afectaciones en etapa temprana de la enfermedad, por lo que se ha incrementado recientemente la cantidad de proyectos que realizan estudios sobre este tipo de imágenes dejando de cierta forma los Rayos X fuera de uso para este tipo de diagnóstico.

La evaluación de diversas redes CNN permite que este estudio sea comparable en resultados con otras investigaciones sobre el tema de estudio y entrega una perspectiva sobre los alcances obtenidos y las áreas de oportunidad.

Uno de los mayores retos que se tuvo en este trabajo fue la falta de datos, debido a lo reciente de la enfermedad, no existen repositorios públicos que permitan generar un modelo entrenado con la robustez que se desea, esto ira cambiando en un futuro ya que en tiempos recientes se ha incrementado la publicación de conjuntos de datos en su mayoría de series tomográficas, ademas en este trabajo se tuvo la aportación de un hospital de puebla que permitió el uso de datos para la alimentación de la red neuronal

y se esta tratando de realizar los tramites de autorización para poder hacer público el conjunto de datos.

Recientemente se ha iniciado el análisis de series topográficas de tórax representadas en 3 dimensiones, esto plantea un reto respecto al estudio y recursos necesarios para obtener diagnósticos aun mas precisos que ademas por el tipo de visualización ayudan aun mas al personal médico. Esta metodología presenta un reto respecto a los recursos de Hardware que requiere para su procesamiento, entrenamiento y clasificación, pero los recientes equipos de computo permiten que sea una opción viable en un tiempo cercano. Se pretende realizar un estudio posterior a este trabajo donde se maneje este tipo de series a partir del conocimiento adquirido.

Un aspecto importante y que tiene cierto valor de análisis para el personal medico es la detección de regiones que han sido afectadas por el virus, por lo que posterior a la clasificación de pacientes enfermos se ve la oportunidad de detectar las áreas afectadas por la neumonía, esto podría beneficiar al diagnostico y tratamiento adecuado para los pacientes. Debido a los limites de tiempo fue imposible realizar un estudio profundo de este tipo de segmentación de lesiones pulmonares pero es un área de oportunidad para trabajos posteriores.

El uso de archivos de datos DICOM, siempre y cuando tengan información complementaria, permite tener ademas de las imágenes un conjunto de datos adicional por ejemplo, sexo, edad, peso, convalecencias, síntomas, diagnostico, por mencionar algunos, que nos permitirían realizar análisis y ensayos mas complejos que pudieran generar predicciones o clasificaciones basándose en algo mas que las imágenes, por ejemplo determinar que un paciente requiere de una atención mas temprana debido a que la correlación que existe entre COVID-19 y obesidad podría generar una aceleración de la neumonía en los pulmones. Con este tipo de análisis se podría tener un sistema mucho muy robusto que auxilie en las labores de los sistemas de salud tratando de prevenir en lo posible un deceso y ademas dirigiendo los tratamientos a personas que se detecte su vida peligrar.

Un elemento importante en el uso de aprendizaje profundo para el tratamiento de imágenes médicas es el preprocesamiento de éstas, en este trabajo se realizo un preprocesamiento relativamente sencillo donde se realizo el ROI (Area de Interes) , posteriormente se realizo la aplicación de el algoritmo de ecualización de histograma CLAHE El objetivo de ayudar al personal médico por medio de una herramienta de aprendizaje profundo que clasifique imágenes para descartar los casos en los que no este presente la enfermedad tiene un impacto en la capacidad de los sistemas de salud, debido a que ayuda a descargar parte del labor de análisis que realizan los médicos, es importante recalcar que como en la mayoría de los clasificadores automáticos, especialmente en áreas tan delicadas como esta, se requiere siempre del análisis final del personal calificado y no debe de tomarse como un sistema completamente autónomo en sus decisiones.

El uso de entornos de desarrollo como Tensorflow permite el desarrollo relativamente rápido de aplicaciones que utilicen el aprendizaje profundo, brinda un conjunto de herramientas de optimización de los resultados y aceleración de los entrenamientos. Ademas provee de mecanismos para la evaluación de estos y así garantizar en cierto grado la fiabilidad del modelo. El uso de este con el lenguaje de programación Python permite una implementación ágil aun para personas ajenas al área de la computación. Como la mayoría de los entornos de desarrollo Tensorflow evoluciona constantemente promoviendo la simplicidad en su utilización y añadiendo nuevas herramientas que

perfeccionen los resultados.

Es importante hacer notar que la enfermedad COVID-19 esta en constante evolución, existen condiciones que modifican y mutan al virus creando diversas cepas, las cuales causan diferentes síntomas, hasta el momento todas las neumonías que se presentan por esta causa siempre tienen las mismas características aunque algunas cepas muestran menores afectaciones al sistema pulmonar, aun así hay que ser cautelosos y continuar con análisis de los conjuntos de datos que puedan aportar mayor información sobre el tema. Actualmente la detección de las variantes en los pacientes no se realiza de forma habitual y habrá que esperar a que los sistemas de salud determinen si es realmente necesario este dato para aportar información relevante a la condición de los pacientes.

Bibliografía

- Azghandi, Marjan y Mohammad Amin Kerachian (2020). *Detection of novel coronavirus (SARS-CoV-2) RNA in peripheral blood specimens*. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7605324/>.
- Chen, Nanshan y col. (2020). *Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study*. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7135076/>.
- Chung, Michael y col. (2020). «CT Imaging Features of 2019 Novel Coronavirus (2019-nCoV)». En: *Radiology* 295.1. PMID: 32017661, págs. 202-207. DOI: [10.1148/radiol.2020200230](https://doi.org/10.1148/radiol.2020200230). eprint: <https://doi.org/10.1148/radiol.2020200230>. URL: <https://doi.org/10.1148/radiol.2020200230>.
- Coronavirus disease (COVID-19)*. URL: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>.
- Esteva, Andre y col. (2017). «Dermatologist-level classification of skin cancer with deep neural networks». En: *Nature* 542.7639, 115–118. DOI: [10.1038/nature21056](https://doi.org/10.1038/nature21056).
- Goodfellow, Ian y col. (2020). «Generative adversarial networks». En: *Communications of the ACM* 63.11, 139–144. DOI: [10.1145/3422622](https://doi.org/10.1145/3422622).
- Gottesman, Omer y col. (2019). «Guidelines for reinforcement learning in healthcare». En: *Nature Medicine* 25.1, 16–18. DOI: [10.1038/s41591-018-0310-5](https://doi.org/10.1038/s41591-018-0310-5).
- Gulshan, Varun y col. (2019). «Performance of a Deep-Learning Algorithm vs Manual Grading for Detecting Diabetic Retinopathy in India». En: *JAMA Ophthalmology* 137.9, pág. 987. DOI: [10.1001/jamaophthalmol.2019.2004](https://doi.org/10.1001/jamaophthalmol.2019.2004).
- He, Kaiming y col. (2016). «Deep Residual Learning for Image Recognition». En: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. DOI: [10.1109/cvpr.2016.90](https://doi.org/10.1109/cvpr.2016.90).
- Joulin, Armand y col. (2015). *Learning Visual Features from Large Weakly Supervised Data*. arXiv: [1511.02251](https://arxiv.org/abs/1511.02251) [cs.CV].
- Kudo, Akira y col. (2019). «Virtual Thin Slice: 3D Conditional GAN-based Super-Resolution for CT Slice Interval». En: *Machine Learning for Medical Image Reconstruction Lecture Notes in Computer Science*, 91–100. DOI: [10.1007/978-3-030-33843-5_9](https://doi.org/10.1007/978-3-030-33843-5_9).
- Lakhani, Paras y Baskaran Sundaram (2017). «Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks». En: *Radiology* 284.2, 574–582. DOI: [10.1148/radiol.2017162326](https://doi.org/10.1148/radiol.2017162326).
- Maguolo, Gianluca y Loris Nanni (2021). «A critic evaluation of methods for COVID-19 automatic detection from X-ray images». En: *Information Fusion* 76, 1–7. DOI: [10.1016/j.inffus.2021.04.008](https://doi.org/10.1016/j.inffus.2021.04.008).
- Murphy, Kevin P. (2013). *Just the facts101 Machine learning: a probabilistic perspective by Kevin P. Murphy. 1st edition*. Cram101.

- Pereira, Sergio y col. (2017). «On hierarchical brain tumor segmentation in MRI using fully convolutional neural networks: A preliminary study». En: *2017 IEEE 5th Portuguese Meeting on Bioengineering (ENBENG)*. DOI: [10.1109/enbeng.2017.7889452](https://doi.org/10.1109/enbeng.2017.7889452).
- Perez, Luis y Jason Wang (2017). «The Effectiveness of Data Augmentation in Image Classification using Deep Learning». En: *CoRR abs/1712.04621*. arXiv: [1712.04621](https://arxiv.org/abs/1712.04621). URL: <http://arxiv.org/abs/1712.04621>.
- Rahimzadeh, Mohammad, Abolfazl Attar y Mohammad Sakhaei (2020). «A Fully Automated Deep Learning-based Network For Detecting COVID-19 from a New And Large Lung CT Scan Dataset». En: DOI: [10.20944/preprints202006.0031.v1](https://doi.org/10.20944/preprints202006.0031.v1).
- Rawat, Waseem y Zenghui Wang (2017). «Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review». En: *Neural Computation* 29.9, 2352–2449. DOI: [10.1162/neco_a_00990](https://doi.org/10.1162/neco_a_00990).
- Redmon, Joseph y col. (2016). «You Only Look Once: Unified, Real-Time Object Detection». En: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. DOI: [10.1109/cvpr.2016.91](https://doi.org/10.1109/cvpr.2016.91).
- Tartaglione, Enzo y col. (2020). «Unveiling COVID-19 from CHEST X-Ray with Deep Learning: A Hurdles Race with Small Data». En: *International Journal of Environmental Research and Public Health* 17.18, pág. 6933. DOI: [10.3390/ijerph17186933](https://doi.org/10.3390/ijerph17186933).
- Valsesia, Diego, Giulia Fracastoro y Enrico Magli (2019). «Image Denoising with Graph-Convolutional Neural Networks». En: *2019 IEEE International Conference on Image Processing (ICIP)*. DOI: [10.1109/icip.2019.8803367](https://doi.org/10.1109/icip.2019.8803367).
- Wang, Hongkai y col. (2017). «Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from 18F-FDG PET/CT images». En: *EJNMMI Research* 7.1. DOI: [10.1186/s13550-017-0260-9](https://doi.org/10.1186/s13550-017-0260-9).
- Yadav, Samir S. y Shivajirao M. Jadhav (2019). «Deep convolutional neural network based medical image classification for disease diagnosis». En: *Journal of Big Data* 6.1. DOI: [10.1186/s40537-019-0276-2](https://doi.org/10.1186/s40537-019-0276-2).
- Zhao, Jinyu y col. (2020). «COVID-CT-Dataset: a CT scan dataset about COVID-19». En: *arXiv preprint arXiv:2003.13865*.
- Zhu, Jun-Yan y col. (2017). «Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks». En: *2017 IEEE International Conference on Computer Vision (ICCV)*. DOI: [10.1109/iccv.2017.244](https://doi.org/10.1109/iccv.2017.244).