

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE MEDICINA
DIVISIÓN DE ESTUDIOS DE POSGRADO

INSTITUTO MEXICANO DEL SEGURO SOCIAL
CENTRO MÉDICO NACIONAL LA RAZA
UMAE HOSPITAL DE ESPECIALIDADES
"DR. ANTONIO FRAGA MOURET"

TÍTULO
DESARROLLO DE UN MODELO DE INTELIGENCIA ARTIFICIAL
CON MÉTODOS DE APRENDIZAJE MÁQUINA PARA EL
DIAGNÓSTICO DE NEUMONÍA EN PACIENTES POR COVID-19
UTILIZANDO RADIOGRAFÍAS DE TÓRAX.

TESIS
PARA OBTENER EL GRADO DE ESPECIALISTA EN
MEDICINA CRÍTICA

PRESENTA
Dr. Tomás Abraham Castillo Hernández

ASESORES
Dra. Nancy Allin Canedo Castillo.
Dr. Alejandro Esquivel Chávez.

CIUDAD DE MEXICO MAYO DE 2022





Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Dr. Jesús Arenas Osuna
Jefe de la División de Educación en Salud
UMAЕ Hospital de Especialidades “Dr. Antônio Fraga Mouret”
Centro Médico Nacional La Raza
Instituto Mexicano del Seguro Social

Dr. Alejandro Esquivel Chávez
Profesor Titular del Curso de Especialización en Medicina Crítica
Unidad de Cuidados Intensivos
UMAЕ Hospital de Especialidades “Dr. Antonio Fraga Mouret”
Centro Médico Nacional La Raza
Instituto Mexicano del Seguro Social

Dra. Nancy Allin Canedo Castillo
Médico adscrito a
Unidad de Cuidados Intensivos
UMAЕ Hospital de Especialidades “Dr. Antonio Fraga Mouret”
Centro Médico Nacional La Raza
Instituto Mexicano del Seguro Social

Residente del Curso de Especialización en Medicina Crítica
Unidad de Cuidados Intensivos
UMAЕ Hospital de Especialidades “Dr. Antonio Fraga Mouret”
Centro Médico Nacional La Raza
Instituto Mexicano del Seguro Social

Número de registro:
R-2022-3501-004

ÍNDICE

RESUMEN	3
ABSTRACT.....	4
ANTECEDENTES CIENTÍFICOS	5
Situación Actual de COVID-19.....	6
Machine learning	8
Redes Neuronales Artificiales	8
Utilización de machine learning en el diagnóstico de neumonía por COVID-19.	10
MATERIAL Y MÉTODO.....	15
Entrenamiento de la red.....	16
ANÁLISIS ESTADÍSTICO	18
RESULTADOS	19
Desempeño	¡Error! Marcador no definido.
TABLAS Y GRAFICOS.....	21
DISCUSIÓN	25
CONCLUSIONES.....	27
REFERENCIAS BIBLIOGRAFICAS	28
ANEXOS	31
Tabla 1. Características de la población según frecuencia y porcentaje.....	21
Tabla 2. Estadísticas descriptivas de variables escalares de acuerdo con el diagnóstico	22
Ilustración 1. Curva ROC para Modelo Entrenado.....	23
Ilustración 2. Curva ROC para Médico Imagenología	24

RESUMEN

DESARROLLO DE UN MODELO DE INTELIGENCIA ARTIFICIAL CON MÉTODOS DE APRENDIZAJE MÁQUINA PARA EL DIAGNÓSTICO DE NEUMONÍA EN PACIENTES CON COVID-19 UTILIZANDO RADIOGRAFÍAS DE TÓRAX.

Introducción: La enfermedad COVID-19 tiene una amplia gama de presentaciones clínicas, desde cuadros asintomáticos hasta falla respiratoria y muerte. Actualmente, la única forma de detectar casos asintomáticos de COVID-19 es mediante programas de cribado masivo; los cuales son poco viables en la mayoría de los países debido a la falta de infraestructura y a la cantidad de recursos necesarios para llevarse a cabo. **Objetivos:** Crear un modelo de inteligencia artificial a partir de radiografías de tórax utilizando modelos de aprendizaje supervisado con alto valor predictivo positivo para identificar neumonía por COVID-19. **Material y métodos:** se analizaron 590 imágenes, divididas en 2 grupos, COVID y no COVID. Los datos se dividieron en: 80% del conjunto para entrenamiento y el 20% para pruebas. Las imágenes se normalizaron y se creó una red convolucional desde cero. Una vez entrenada la red se analizó el grupo de pruebas y se calculó el desempeño, posteriormente, la concordancia interobservador de una submuestra con médico radiólogo especialista con experiencia en radiografía simple de tórax y COVID-19. **Resultados:** Se obtuvo un modelo computacional con Sensibilidad 0.99, Especificidad 0.65, VPP 0.85, VPN 0.98, así como Accuracy total: 0.8810 y un AUC 0.86. Comparado con el desempeño obtenido por médico especialista Sensibilidad 0.79, Especificidad 0.88, VPP 0.9, VPN 0.76. así como de Accuracy 0.83 y AUC 0.83, el índice kappa entre el modelo y la PCR fue de 0.72, entre imagenología y PCR 0.66 y entre ambos observadores de -0.132. **Conclusiones:** El modelo desarrollado obtuvo una Precisión media (ACCURACY) del 88% y presenta un mejor desempeño como prueba de pesquisa debido a contar con sensibilidad alta (99%) y valor predictivo negativo alto (98%).

Palabras clave: COVID-19, diagnóstico, radiografía de tórax, inteligencia artificial, neumonía.

ABSTRACT

DEVELOPMENT OF AN ARTIFICIAL INTELLIGENCE MODEL WITH MACHINE LEARNING METHODS FOR THE DIAGNOSIS OF PNEUMONIA IN PATIENTS WITH COVID-19 USING CHEST RADIOGRAPHS.

Introduction: COVID-19 has a wide range of clinical presentations, from asymptomatic to respiratory failure and death. Currently, the only way to detect asymptomatic cases of COVID-19 is through mass screening programs, which are not feasible in most countries due to the lack of infrastructure and the number of resources required to carry them out.

Objectives: To create an artificial intelligence model from chest radiographs using supervised learning models with high positive predictive value to identify COVID-19 pneumonia. **Methods:** 590 images were analyzed, divided into 2 groups, COVID and non-COVID. The data were divided into 80% of the set for training and 20% for testing. The images were normalized, and a convolutional network was created from scratch.

Once the network was trained, the test set was analyzed and the performance was calculated, subsequently, the interobserver concordance of a subsample with a specialist radiologist with experience in simple chest radiography and COVID-19 was calculated.

Results: A computational model with Sensitivity 0.99, Specificity 0.65, PPV 0.85, NPV 0.98, as well as total Accuracy 0.8810 and AUC 0.86 was obtained. Compared to the performance obtained by specialist physician Sensitivity 0.79, Specificity 0.88, PPV 0.9, NPV 0.76. as well as Accuracy 0.83 and AUC 0.83, the kappa index between model and CRP was 0.72, between imaging specialist and CRP 0.66 and between both observers - 0.132. **Conclusions:** The developed model obtained a mean Accuracy of 88% and presents a better performance as a screening test due to having high sensitivity (99%) and high negative predictive value (98%).

Key words: COVID-19, diagnosis, chest radiography, artificial intelligence, pneumonia.

ANTECEDENTES CIENTÍFICOS

La enfermedad provocada por el virus SARS-CoV-2, denominada como COVID-19, tiene una amplia presentación clínica y en algunos casos inespecífica. La infección puede manifestarse como un portador asintomático y en otros casos puede observarse falla respiratoria y muerte¹. Un reporte reciente señaló que las infecciones no documentadas o asintomáticas facilitan la rápida diseminación de COVID-19, por lo que su control es sumamente complejo².

Actualmente, la única forma de detectar casos asintomáticos de COVID-19 es mediante programas de cribado masivo como aquellos aplicados en Corea del Sur, donde la capacidad de prueba alcanzó hasta 15,000 pruebas por día a población general y se promulgó una legislación que otorgó al gobierno autoridad para recopilar datos de teléfonos móviles, tarjetas de crédito y otros datos de quienes dan positivo para reconstruir su paradero reciente y localización de posibles contactos², medidas que son poco viables en la mayoría de los países - especialmente en aquellos en vías de desarrollo - debido a la falta de infraestructura, a la cantidad de recursos humanos y económicos necesarios para que estos programas puedan llevarse a cabo. Por esta razón, es necesaria una herramienta de cribado alternativa con la suficiente robustez para discriminar entre los casos de COVID-19, que podría reducir sustancialmente costos tanto económicos como humanos en el combate de la pandemia actual, especialmente en países donde un cribado generalizado con las técnicas dependientes de tecnología actuales no es una opción.

La creación de un método para diagnosticar neumonía por COVID-19 de bajo costo, de fácil implementación y accesible para la población en general supondría un cambio de paradigma y una forma de cortar la transmisión del virus. Con esto en mente, en el presente protocolo planteamos la creación de un modelo robusto de inteligencia artificial utilizando redes neuronales artificiales para la implementación de un sistema de clasificación a partir de radiografías de tórax que resulte fácil de aplicar, utilizar e implementar.

Situación Actual de COVID-19

En diciembre del 2019, China reportó un brote de síndrome respiratorio agudo severo (SARS, por sus siglas en inglés) causado por un nuevo coronavirus en la provincia de Wuhan, la cual ha sido designada por la Organización Mundial de la Salud (OMS) como enfermedad del coronavirus 2019 (COVID-19)¹. Esta enfermedad se ha extendido mundialmente desde su aparición; hasta el momento existen cerca de 5 millones de personas infectadas y 315,000 fallecidos, y los números incrementan diariamente³.

Actualmente, las recomendaciones realizadas por la OMS para contener y mitigar la pandemia tienen como objetivo el rompimiento de la cadena de transmisión. Por lo tanto, las acciones tomadas hasta el momento incluyen el distanciamiento social, higiene de manos, seguimiento epidemiológico y, recientemente, la aplicación masiva de pruebas diagnósticas para la COVID-19⁴. Sin embargo, esta última recomendación podría resultar infectiva, ya que modelos matemáticos actuales muestran que, si la cadena de transmisión no se rompe, la tasa de infección aumentará exponencialmente y una gran proporción del mundo será infectado, por lo que la cantidad de pruebas realizadas será insuficiente a pesar de que la OMS ha mandado 1.5 millones de pruebas diagnósticas a 120 países diferentes⁵.

La enfermedad tiene una amplia presentación clínica que abarca desde casos asintomáticos hasta falla respiratoria y muerte¹. Debido a esta gran variedad de cuadros, resulta preocupante un reporte reciente en donde se señala que infecciones no documentadas o asintomáticas facilitan la rápida diseminación de la COVID-19⁶. Actualmente, la única forma de detectar casos asintomáticos de la COVID-19 es mediante programas de cribado masivo como aquellos aplicados en Corea del Sur², los cuales son poco viables en la mayoría de los países. Especialmente en aquellos en vías de desarrollo, dada la falta de infraestructura y la cantidad de recursos humanos y económicos necesarios para que estos programas se lleven a cabo. Por esta razón, es necesaria una herramienta de cribado alternativa con la suficiente sensibilidad y especificidad para descartar casos de neumonía la por COVID-19, ya que podría representar una disminución sustancial en costos tanto económicos como humanos en el combate de la pandemia actual.

De acuerdo con el Centro Nacional de Inmunización y Enfermedades Respiratorias (NCIRD por sus siglas en inglés), División de Enfermedades Virales, de los Estados Unidos de Norte América las pruebas de SARS-CoV-2 pueden incorporarse como parte de un enfoque integral para reducir la transmisión. La detección de síntomas, las pruebas y el rastreo de contactos son estrategias para identificar a las personas infectadas con el SARS-CoV-2, de modo que se puedan tomar medidas para frenar y detener la propagación del virus.⁷

Junto con el avance en el diagnóstico médico, los enfoques basados en la detección de ácidos nucleicos se han convertido en una tecnología rápida y confiable para la detección de virus. Las pruebas de amplificación de ácido nucleico detectan uno o más genes de ácido ribonucleico viral (ARN) e indican una infección actual o una infección reciente, pero, debido a la detección prolongada de ARN viral, no siempre son evidencia directa de la presencia de virus capaces de replicarse o transmitirse a otros. La mayoría de las pruebas de amplificación de ácidos nucleicos deben procesarse en un laboratorio y el tiempo para obtener los resultados puede variar (~ 1 a 3 días), algunas pruebas consideradas "rápidas" cuentan con resultados disponibles en aproximadamente 15 a 45 minutos.⁷ La mayoría de las pruebas de amplificación de ácidos nucleicos producen resultados cualitativos. Entre las pruebas de ADN, el método de reacción en cadena de la polimerasa (PCR) se considera el "estándar de oro" para la detección de muchos virus y se caracteriza por una detección rápida, alta sensibilidad y especificidad.⁸

Uno de los principales problemas de la RT-PCR en tiempo real es el riesgo de resultados falsos negativos y falsos positivos. Se ha informado que muchos casos "sospechosos" con características clínicas típicas de COVID-19 y tomografía computarizada (TC) específica no han sido diagnosticados.⁸ Por lo tanto, un resultado negativo no excluye la posibilidad de COVID-19 y no debe usarse como el único criterio para el manejo del paciente o decisiones de tratamiento. La combinación de RT-PCR en tiempo real, características clínicas e interpretación de estudio de imagen parece facilitar el manejo de los brotes de SARS-CoV-2.

Machine learning

El aprendizaje automático o machine learning por su nombre en inglés, es una rama de las ciencias de la computación en la cual un algoritmo computacional puede aprender a realizar una tarea específica sin la necesidad de ser específicamente programado para esta tarea. Estos algoritmos pueden resolver tareas que comúnmente se clasifican en dos tipos: problemas de regresión y problemas de clasificación⁹; en el caso de este proyecto, nos enfocaremos en los problemas de clasificación⁹.

Redes Neuronales Artificiales

Dentro del campo de los problemas de clasificación, las redes neuronales artificiales han mostrado ser una herramienta poderosa para la creación de modelos de inteligencia artificial en donde los datos son complejos (imágenes, voz, etc). Este método de *machine learning* toma su nombre de la organización intrínseca que ocurre en el sistema nervioso en donde neuronas interconectadas intercambian información; organización que normalmente incluye neuronas de entrada—a las cuales se les alimentan los datos de interés—neuronas de procesamiento y neuronas de salida—las cuáles nos proveerán del resultado de la clasificación⁹. La red neuronal más sencilla es denominada perceptrón y se encuentra ilustrada en el anexo 1.

En un perceptrón cada uno de los datos $X_0, X_1, X_2 \dots X(n)$ representa un dato de entrada o variable independiente de nuestro modelo de predicción, los cuales son multiplicados de manera independiente por un “peso” (weight) relacionado con la neurona de salida (los pesos se encuentran representados por $W_0, W_1, W_2 \dots W(n)$); el peso determinará la fuerza de la conexión o la influencia que tendrá la neurona de entrada con la neurona de salida.

Por otro lado, con la letra b se representa el ajuste (*bias*, por su nombre en inglés), el cual permite modificar la conectividad entre neuronas hacia arriba o hacia abajo. En resumen, cada uno de los datos de entrada (x) es multiplicado por el peso correspondiente (w), sumados a los demás datos y al ajuste (b); una vez realizada esta sumatoria, se pasa por una función de activación (f) y genera un *output* o salida de la red neuronal⁷.

Las funciones de activación determinarán el comportamiento final de las redes neuronales artificiales, convirtiendo los datos de entrada que convergen en un solo nodo a un solo dato de salida. Por lo tanto, dependiendo de la función de activación, la neurona podrá o no ser activada; permitiéndole utilizar los datos de entrada de manera no lineal y, de esta forma, brindándole la capacidad de aprender de datos complejos. Existen diversas funciones de activación donde las más comunes son la activación por umbral, activación sigmoidea, hiperbólica tangente y rectificadora de unidades lineales⁹. Asimismo, es importante mencionar que, en el caso de que una red neuronal tenga muchas capas, este dato de salida servirá como dato de entrada para la próxima capa de neuronas⁹.

Para que una red neuronal artificial pueda aprender, requiere retroalimentación de parte del programador sobre la predicción que realizó, es decir, que le dé datos sobre lo correcto o incorrecto en su predicción; una vez realizada una predicción, basándose en la diferencia entre el valor real y el valor predicho, se calcula un valor de error denominado función de costo (*cost function*). A continuación, esta función de costo es analizada para modificar los valores de peso (w) y de la función de activación (f) con el objetivo de mejorar el desempeño del modelo al hacer que el valor de la función de costo sea lo más baja posible, haciendo a las predicciones del modelo más aproximadas a la realidad. Este proceso se repite muchas veces en aras de encontrar la mejor combinación de pesos para obtener la función de costo más baja, proceso que se denomina propagación hacia atrás (*back-propagation*)⁹.

Existen diversas formas de implementar un algoritmo de propagación hacia atrás, siendo la utilización del gradiente la más efectiva para datos complejos. El gradiente es la extensión matemática de la derivada en el cálculo multivariable y su interpretación en un campo tridimensional se refiere a la dirección hacia la cual la función crece más rápidamente dado un punto en el campo tridimensional. El concepto de gradiente no se limita a tres dimensiones y se puede utilizar con N cantidad de estas, lo que es ideal para optimizar la función de propagación hacia atrás si se toma cada peso como la dimensión de un vector; el negativo del gradiente de este vector n -dimensional nos indicará hacia donde se debe modificar cada uno de los pesos para hacer que la función de costo tienda al mínimo⁹.

Casi todo problema puede ser reducido a una función matemática dada, sin embargo, es imposible para un ser humano identificar todas las dimensiones de un problema complejo para modelar una función específica. En este sentido, las redes neuronales artificiales funcionan como un “aproximador universal de funciones”, por lo que puede modelar la función desconocida, aun cuando no sea obvia para un ser humano⁹. Siendo así, las redes neuronales artificiales formarán parte del método base de la presente propuesta.

[Utilización de machine learning en el diagnóstico de neumonía por COVID-19.](#)

Avances recientes en la utilización de machine learning e inteligencia artificial han mostrado ser capaces de diagnosticar enfermedades a través de imágenes (radiografías, fotografías, tomografías) con niveles de confiabilidad similares e inclusive mejores a especialistas médicos¹⁰.

A partir del aviso hecho por la oficina en China de la OMS se aceleró la generación de conocimiento sobre la nueva enfermedad; en los primeros días, las publicaciones se concentraron en instituciones chinas, pero al hacerse más claro el peligro de la extensión global del brote, científicos de todo el mundo comenzaron a trabajar en el tema. Una de las áreas con mayor auge fue el desarrollo de métodos diagnósticos que permitiesen identificar a la población afectada con menor número de recursos y menores tiempos de espera, con esto en mente se empezaron a desarrollar métodos diagnósticos a través del uso de imagen y su análisis por métodos computacionales, como las redes neuronales.

De los primeros métodos alternativos utilizados para el diagnóstico de COVID -19, fue un modelo de redes neuronales artificiales utilizando tomografías computadas fue capaz de lograr una clasificación de casos de COVID-19 con una precisión total de 89.5% con una especificidad de 88% y una sensibilidad del 87%. Aunque el algoritmo muestra resultados interesantes, la cantidad de tomografías utilizadas es demasiado baja (solamente 79 casos de COVID-19) y la utilización de tomografías computadas incluye una alta dosis de radiación que no sería pertinente para un método de cribado universal¹¹.

Posteriormente, se ha impulsado el desarrollo de otros métodos de diagnóstico asistido por computadora (CAD), principalmente aquellos basados en radiografías de tórax (CXR), ya que son una de las alternativas más baratas para diagnosticar etapas tempranas de la enfermedad COVID-19 en comparación con otras alternativas como pudieran ser la reacción en cadena de la polimerasa (PCR), tomografía computarizada (CT), entre otras. Asimismo, el uso de métodos CAD resulta novedoso ya que podría lograr exitosamente la diferenciación de COVID-19 a partir de radiografía de tórax con otras infecciones virales, algo que ha resultado difícil para los radiólogos debido a que el patrón observado es muy similar entre los dos tipos de infección^{12,13}; no obstante, los modelos de inteligencia artificial basados en redes neuronales artificiales han mostrado ser muy efectivos para la clasificación de imágenes radiológicas, aun cuando los patrones no son claros para los seres humanos¹⁰. Con este fin, se han propuesto algunos trabajos para diagnosticar COVID-19 mediante el uso de métodos basados en CXR.

Entre los primeros ensayos, se presentó un reporte empírico utilizando radiografías de tórax y redes neuronales artificiales mostró resultados prometedores al lograr una precisión en la clasificación del 92% de las radiografías con una sensibilidad del 100% y especificidad del 80%¹⁴. Posteriormente, se han propuesto modificaciones a la estructura VGG16 para lograr hacer esta clasificación diagnóstica; trabajo realizado por Siaula C. et al. que obtuvo un rendimiento prometedor para la clasificación de imágenes CXR en el diagnóstico de COVID-19.¹⁵

La VGG16 es un modelo de red neuronal convolucional (CNN) propuesto por Karen Simonyan y Andrew Zisserman en la Universidad de Oxford presentado durante el desafío The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) en 2014, una competencia anual que evaluó algoritmos para la clasificación de imágenes (y detección de objetos) a gran escala. Fue el ganador del desafío de detección de ImageNet en 2014. VGG16 es uno de los modelos más utilizados para el diagnóstico de enfermedades utilizando imágenes médicas como rayos X o resonancia magnética debido a su principal función como herramienta para el reconocimiento o clasificación de imágenes.¹⁶

Por otra parte, se han buscado desarrollar y validar nuevos sistemas de diagnóstico asistido por computadora para la clasificación de casos entre neumonía por COVID-19, neumonía no COVID-19 e imágenes sanas en la radiografía de tórax. Nishio M. y colaboradores trabajaron a partir de dos conjuntos de datos públicos, donde se obtuvieron un total de 1248 imágenes de CXR, que incluían 215, 533 y 500 imágenes de CXR de pacientes con neumonía por COVID-19, pacientes con neumonía sin COVID-19 y muestras sanas, respectivamente. Las 1248 imágenes de CXR se dividieron en 998, 125 y 125 imágenes para entrenamiento, validación y conjuntos de prueba, respectivamente. Para la normalización de imágenes, las imágenes de CXR se dividieron por 255, y los valores de píxel de ellas oscilaron entre 0 y 1. El sistema CXDx propuesto utilizó VGG16, un modelo pre entrenado para análisis de imágenes modificado con técnicas de aumento de datos. La evaluación del rendimiento se realizó utilizando la precisión de clasificación en 3 categorías del conjunto de prueba (125 imágenes de CXR). La precisión media (ACCURACY) del método propuesto fue del 83,7 % y la sensibilidad media de la neumonía por COVID-19 fue del 90,9 %.¹⁷

Además, Elaziz et al. realizaron un estudio experimental transversal comparativo, buscando proponer un método de clasificación para imagen de rayos X de tórax de COVID-19 que depende de la combinación de la fuerza de dos técnicas.

Primero, un nuevo descriptor de imagen, Fractional Multichannel Exponent Moments (FrMEMs), encargado de extraer características de las imágenes de rayos X y, en segundo lugar, una técnica de selección de características modificada basada en Manta-Ray Foraging Optimization y evolución diferencial (MRFODE) para mejorar la capacidad de encontrar las características relevantes de esas características extraídas. Las características extraídas se dividen en dos, conjuntos de entrenamiento y prueba, que representan el 80 % y el 20 % respectivamente del número total de imágenes. Su enfoque propuesto logró altos rendimientos, índices de precisión del 96,09 % y 98,09 % para el primer y segundo conjunto de datos, respectivamente.¹⁸

Por su parte, Apostolopoulos et al. Utilizaron un conjunto de datos de imágenes de rayos X de pacientes con neumonía bacteriana, enfermedad confirmada de COVID-19 e

imágenes normales para la detección automática de la enfermedad por coronavirus con el objetivo de evaluar el rendimiento de las arquitecturas de redes neuronales convolucionales (CNN) de última generación propuestas en los últimos años para la clasificación de imágenes médicas. Procesaron una colección de 1427 rayos X torácicos para entrenar y probar las CNN. Los resultados obtenidos por dicho grupo sugieren que el aprendizaje profundo con imágenes de rayos X puede extraer biomarcadores significativos relacionados con la enfermedad de Covid-19, mientras que la mejor precisión, sensibilidad y especificidad obtenidas es del 96,78 %, 98,66 % y 96,46 %, respectivamente.¹⁹

A medida que avanza la pandemia y se van perfeccionando técnicas de identificación de imágenes asistidos por computadora para el diagnóstico de enfermedades, distintos grupos presentan trabajos con el objetivo de validar nuevos modelos de redes neuronales; tal es el caso de Abbas A, quien publica un trabajo en el cual se valida una CNN llamada DeTraC, para la clasificación de imágenes de rayos X de tórax COVID-19 en el que se utilizaron 2 bases de datos, una de radiografía de tórax normal de la sociedad japonesa de tecnología radiológica (JSRT por sus siglas en inglés) y otra de radiografía de tórax del “COVID-19 open image data collection” la cual incluyó 105 muestras de COVID-19 y 11 muestras de y SARS. Utilizando técnicas de amplificación de datos se generó un total de 1764 muestras, divididas en 2 grupos: 70% para entrenamiento del modelo y 30% para la evaluación del desempeño del modelo. Para la evaluación del desempeño, se tomaron valores de Precisión (ACC), Especificidad (SP) y Sensibilidad (SN). DeTraC se validó con diferentes modelos CNN pre entrenados, donde la mayor precisión la obtuvo VGG19 (una versión mejorada del popular VGG16) en DeTraC con resultados de precisión del 97,35 %, sensibilidad del 98,23 % y especificidad del 96,34%. Y un AUC: 0.96549²⁰.

El aprendizaje por transferencia es una estrategia preferible para entrenar las CNN profundas. El aprendizaje por transferencia es una técnica de aprendizaje automático que consiste en reutilizar una red neuronal entrenada y existente, desarrollada para una tarea, como base para otra tarea. Transferir el aprendizaje elimina la necesidad de

grandes datos de entrenamiento para una tarea, ya que las características básicas requeridas para entrenar un modelo se importan de la tarea anterior.²¹

Para demostrar la utilidad del aprendizaje por transferencia, KC comparó varios modelos de aprendizaje profundo de última generación; el entrenamiento, la validación y la prueba de estos modelos se realizaron en imágenes de rayos X de tórax (CXR) pertenecientes a cinco clases distintas, contenidas en un conjunto de datos de ImageNet, con un total de 760 imágenes. Estos modelos lograron una precisión de validación del 100 % en todo el conjunto de datos, clasificando de manera efectiva imágenes saludables de neumonía bacteriana, neumonía viral, COVID-19 y SARS basados en imágenes de rayos X de tórax. Con esto además ImageNet demostró ser una fuente confiable para transferir características aprendidas en clasificación de enfermedades por medio de radiografías de tórax.²¹

Debido a que la mayoría de los artículos que tratan sobre diferentes protocolos para el diagnóstico automático de COVID-19 a partir de imágenes de rayos X reportan desempeños muy altos, Maguolo et al. decidieron comparar y evaluar el desempeño de estos protocolos diseñados al cambiar a negro el centro de la exploración de rayos X y entrenar los clasificadores solo en la parte exterior de las imágenes. Haciendo esto pudieron eliminar los pulmones de las imágenes y aun así obtener resultados de rendimiento similares. Por lo tanto, se dedujo que varios protocolos de prueba para el reconocimiento no son legítimos y que las redes neuronales están aprendiendo patrones en el conjunto de datos que no están correlacionados con la presencia o no de COVID-19²². Con esto en mente, en este trabajo proponemos un nuevo método de clasificación que deje de lado estos errores y pueda lograr el objetivo de clasificación.

MATERIAL Y MÉTODO

Se llevó a cabo un estudio no experimental, transversal, retrospectivo, analítico y unicéntrico de marzo de 2019 a febrero de 2022. Se recolectó la información de forma retrospectiva a través del expediente del paciente y el sistema electrónico INFOSALUD Compañía Mexicana de Radiología (versión 0.17.13.0) para visualización de imágenes de rayos x, proyección anteroposterior de tórax de pacientes hospitalizados entre el 1º de marzo de 2020 al 30 de octubre del 2020 con diagnóstico de neumonía por COVID-19. Así como, se localizó el expediente electrónico de pacientes con radiografía de tórax sin presencia de enfermedad pulmonar y pacientes con los siguientes diagnósticos en los últimos 2 años de Neumonía bacteriana, Neumonía por COVID-19 + Infección Bacteriana Agregada y otras causas de Neumonía (No Bacteriana, No COVID-19).

Se incluyen 705 radiografías de tórax, las cuales se dividieron en dos grupos COVID y No COVID, para el primer grupo en posición anteroposterior con los metadatos asociados: ID del paciente, diagnóstico etiológico del paciente por algún tipo de prueba de biología molecular o microbiología, sexo, edad, fecha de radiografía. Para el segundo grupo (llamado también control) compuesto por imágenes de radiografías en posición anteroposterior de tórax que cumplieron con alguno de los siguientes criterios: estado de ventilación mecánica, paciente con diagnóstico de neumonía por una causa diferente a COVID-19, paciente con enfermedad pulmonar obstructiva crónica, paciente con edema agudo pulmonar de causa cardiogénica, los metadatos anteriormente mencionados, además de fecha de inicio de hospitalización, supervivencia, antecedente de diabetes, hipertensión arterial, insuficiencia renal, insuficiencia cardíaca crónica, temperatura, frecuencia cardíaca, Score SOFA y Dímero D.

Se excluyeron las radiografías en las que no se pudo identificar los siguientes metadatos: ID del paciente, diagnóstico etiológico por algún tipo de prueba de biología molecular o microbiología, sexo, edad, fecha de la toma de la radiografía; respecto a la radiografía: mala técnica, la no existencia de relación temporal con el padecimiento del paciente, malformación física aparente en el tórax y dextrocardia. Se eliminaron todas aquellas en las que no fue posible realizar el pre-procesamiento de la radiografía, captura correcta de alguno de los datos mínimos requerido. Por lo anterior las imágenes utilizadas fueron

catalogadas como 'covid' y 'control' (renombrada 'no-covid') y se codificaron como 'covid' = 0 y 'no-covid' = 1. Mientras que las imágenes procedentes de infecciones polimicrobianas, bacterianas o fúngicas no se tomaron en cuenta. Para un total de 590 imágenes analizadas.

Todas las imágenes se recortaron desde el centro con un ancho y largo de 800 px. Estas imágenes se cambiaron de tamaño a 256 x 256 px; luego se convirtieron a escala de grises (de 0 a 255 valores de intensidad) y finalmente se aplicó contraste con valor alfa de 0.5.

Posteriormente, se preparó una submuestra de las 590 imágenes para análisis por médico especialista en Imagenología diagnóstica y terapéutica con experiencia en radiografía simple de tórax y covid-19; con un total de 100 imágenes: 50 covid y 50 no-covid.

Entrenamiento de la red

En cuanto a la etapa de entrenamiento, se dividieron los datos en 80% del conjunto para entrenamiento y 20% para pruebas. Además, las imágenes se normalizaron (de 0 a 1 valores de intensidad) y se creó una red convolucional desde cero con capas:

- Capa convolucional 2D, 32 filtros de 3 x 3, activación ReLU
- Capa max pooling 2D, de 2 x 2
- Capa convolucional 2D, 64 filtros de 3 x 3, activación ReLU
- Capa max pooling 2D, de 2 x 2
- Capa convolucional 2D, 64 filtros de 3 x 3, activación ReLU
- Capa max pooling 2D, de 2 x 2
- Operador flatten
- Capa densa de 64 neuronas tipo ReLU
- Capa densa de 2 neuronas (salida)

Posteriormente, se entrenó por 20 épocas con el optimizador Adams, medida de pérdida Cross Entropy, métrica Accuracy; todos los valores por defecto de Keras. Una vez

entrenada la red se analizó el grupo de pruebas y de acuerdo con el resultado obtenido por el modelo se realizó una tabla 2x2 para cálculo de desempeño.

ANÁLISIS ESTADÍSTICO

Como medidas de desempeño del modelo, se incluyeron exactitud, precisión, sensibilidad, especificidad (incluyendo VPP, VPN y LRs) y F1-score. La utilidad del desarrollo del modelo de inteligencia artificial se evaluó mediante su capacidad discriminativa (área bajo la curva ROC y su intervalo de confianza del 95%). La confiabilidad de la prueba se calculó con el índice Kappa de Cohen (κ), que toma el acuerdo relativo observado entre dos jueces o $\text{Pr}(a)$, la probabilidad de que el acuerdo se deba al azar o $\text{Pr}(e)$ y la diferencia observada causada por el grado de acuerdo posible más allá del azar, descrito en la fórmula:

$$\kappa = \frac{\text{Pr}(a) - \text{Pr}(e)}{1 - \text{Pr}(e)}$$

Dando un índice entre 0 y 1 que indicará el grado de acuerdo; siendo menor al acercarse a 0 y casi perfecto al acercarse a 127.

Se efectuó como control de calidad la concordancia interobservador e intraobservador con el cálculo del coeficiente Kappa anteriormente detallado con la ayuda de la médico especialista en radiología Dra. Tania Valverde para verificar la concordancia, la evaluación se realizó a ciegas con respecto al diagnóstico definitivo y a la opinión de los participantes en el presente estudio o de la misma persona, se verificaron tanto la concordancia absoluta como específica, ésta última a los casos diagnosticados como positivos o negativos a neumonía por SARS-Cov-2.

Un valor de $p < 0.05$ se consideró estadísticamente significativo. Para el análisis estadístico se utilizará el paquete Statistical Package for Social Science, SPSS 20.0 para Windows, Chicago III).

RESULTADOS

De la tabla 2x2 creada con los datos del modelo computacional se obtuvo una Sensibilidad de 0.99, Especificidad 0.65, VPP 0.85, VPN 0.98. Asimismo, se obtuvo un Accuracy de entrenamiento: 1.0000 y un Loss de entrenamiento: 0.0023. Por otra parte, en cuanto a la prueba, se obtuvo un Accuracy de prueba de 0.89 y un Loss de prueba de 0.52.

Se probaron todas las imágenes sobre el mismo modelo entrenado y se obtuvo: Accuracy total: 0.88; Loss total de 73.87 (acumulado). Se calculó un índice de Kappa de Cohen entre el modelo computacional y la prueba PCR con un valor de 0.72 (IC 95% 0.92-0.52). Mediante esta representación de los pares (1-especificidad, sensibilidad) obtenidos, curva ROC con AUC 0.86 (Ver Ilustración 1).

Comparando el desempeño del médico especialista se obtuvo Accuracy 0.83, Sensibilidad 0.79, Especificidad 0.88, VPP 0.9, VPN 0.76. Índice Kappa de Cohen entre la medico de imagenología y la prueba PCR de 0.66 (IC 95% 0.85-0.47), curva ROC con AUC de 0.83 (Ver Ilustración 2).

Ambos observadores obtuvieron una estimación del grado de acuerdo con la prueba microbiológica bueno (0.6-0.8).

Una vez analizado el desempeño de ambos observadores, se realizó el cálculo del índice kappa de Cohen para evaluar la concordancia entre ellos, se encontró un 56% de coincidencias, sin embargo, se estimó hasta un 61% de coincidencias atribuibles al azar, con lo que se calculó un resultado del índice kappa de -0.132 concluyendo que existe discrepancia entre ambos observadores.

Se realizó un análisis de frecuencias para identificar la cantidad de datos válidos por variable de tipo nominal. Asimismo, se realizó por variable y sus niveles para describir las características de la población analizada, incluyendo sexo (masculino y femenino), diagnóstico (no COVID y COVID), supervivencia (murió y sobrevivió), antecedente de comorbilidades (diabetes, hipertensión, falla cardiaca crónica, falla renal crónica), datos

clínicos presentados a su ingreso (fiebre, tos, anosmia), desarrollo de lesión renal aguda, así como estado de ventilación mecánica (no intubado e intubado) (Ver Tabla 1).

Posteriormente se realizó un análisis de distribución de la población de cada uno de los niveles de las variables categóricas según el diagnóstico de los pacientes, es decir, si el paciente padecía o no COVID-19 (Ver Tabla 2).

Se obtuvieron un total de 705 imágenes de radiografía de tórax, de las cuales 197 fueron imágenes de paciente sano o control, 393 proceden de pacientes con diagnóstico confirmado de COVID-19, 73 corresponden a infecciones polimicrobianas, 3 corresponden a neumonía por hongos y 40 a neumonía grave de origen bacteriano.

De las imágenes obtenidas el 63% (445) correspondieron a pacientes del sexo masculino y el 37% (261) al sexo femenino, la media de edad de todas las imágenes fue de 55.17 \pm 15.91 años, de todas las imágenes obtenidas el 46% (325) correspondió a pacientes con intubación endotraqueal y el 54% (381) a pacientes sin intubación. La comorbilidad más frecuentemente presentada fue la hipertensión arterial en un 37.1% seguida por diabetes en 35.3% y la mortalidad general fue de 47.9%.

Los pacientes identificados con diagnóstico de COVID-19 tuvieron tiempos de estancia hospitalaria más largos, valores de oxigenación más bajos (SO₂ y PaO₂/FiO₂), niveles de azoados y presiones de la vía aérea más elevadas, sin embargo, puntajes de SOFA score mas bajos.

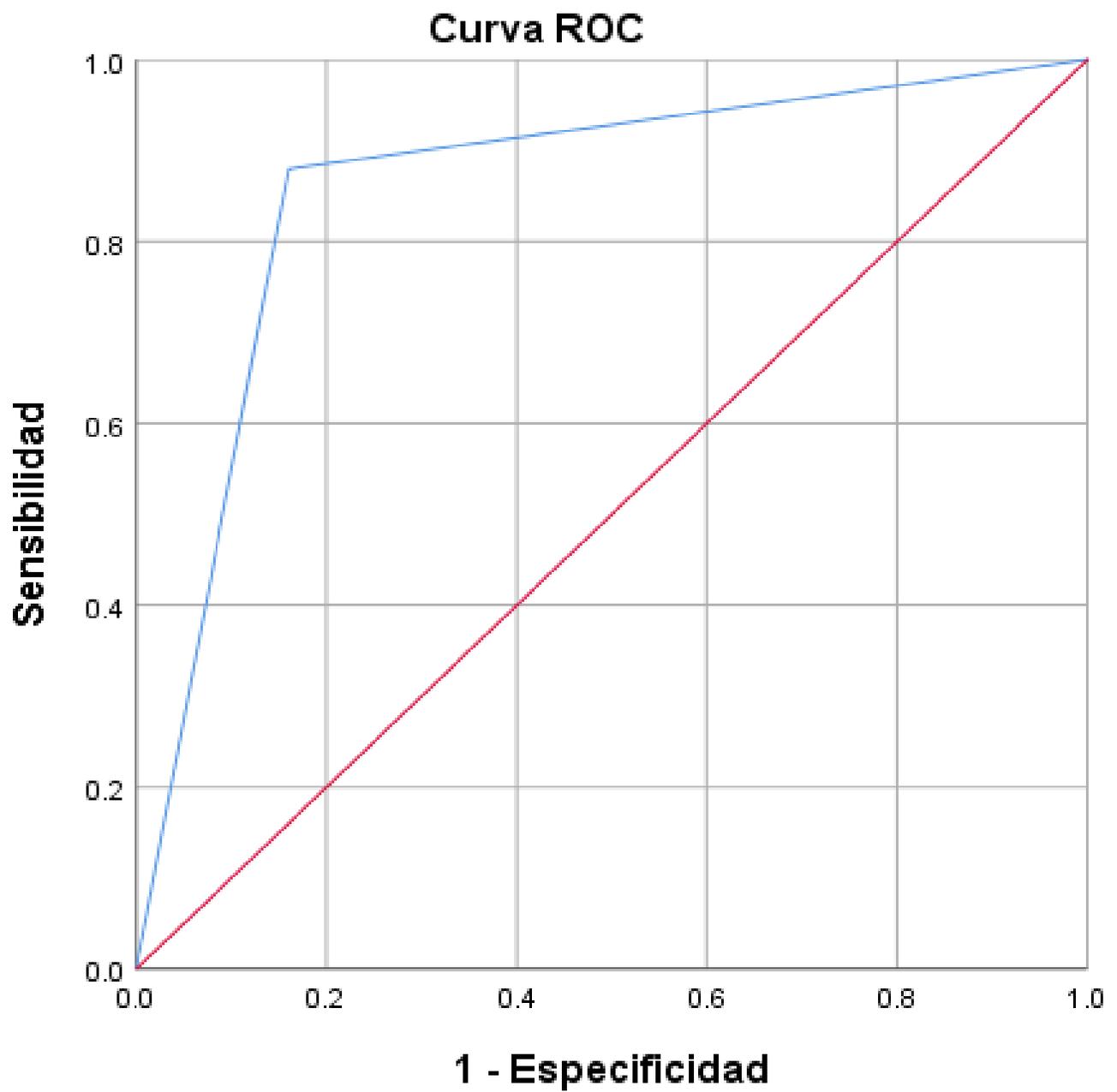
TABLAS Y GRAFICOS

Tabla 1. Características de la población según frecuencia y porcentaje n= 706

	Frecuencia	Porcentaje (%)
Sexo		
Masculino	445	63.0
Femenino	261	37.0
Diagnóstico		
No COVID	240	34.0
COVID	466	66
Supervivencia		
Murió	338	47.9
Sobrevivió	361	51.1
Diabetes		
Sin diabetes	457	64.7
Diabético	249	35.3
Hipertensión		
Sin hipertensión	444	62.9
Hipertenso	262	37.1
Falla cardiaca crónica		
Sin falla cardiaca	598	84.7
Falla cardiaca	102	14.4
Falla renal crónica		
Sin falla renal	540	76.5
Falla renal	161	22.8
Fiebre		
Afebril	319	45.2
Fiebre	249	35.3
Tos		
Sin Tos	212	30.0
Tos	55	7.8
Anosmia		
Sin anosmia	177	25.1
Anosmia	72	10.2
Falla renal aguda		
Sin falla renal aguda	230	32.6
Falla renal aguda	173	24.5
Intubación		
No intubado	325	46.0
Intubado	381	54.0

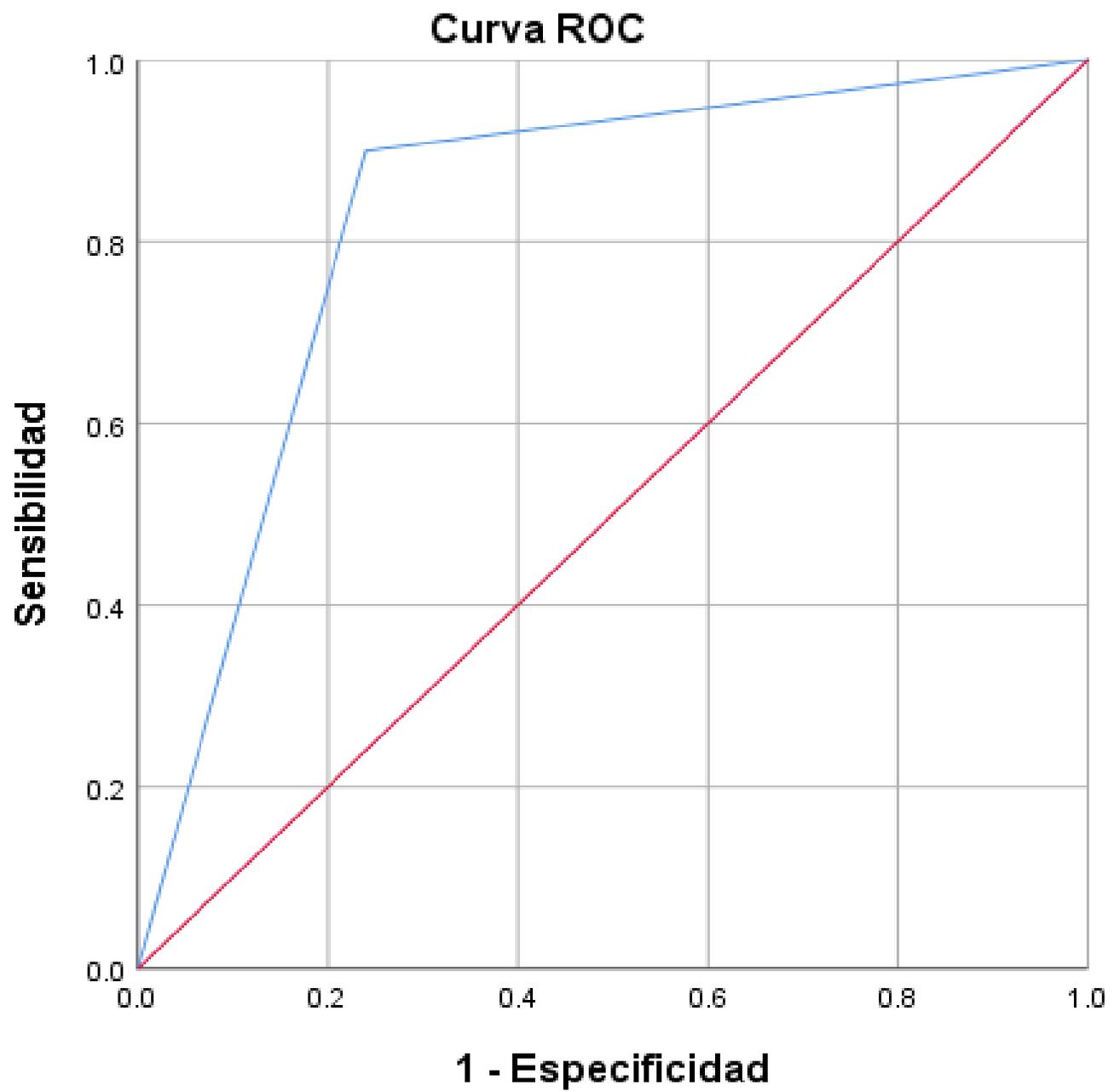
Tabla 2. Estadísticas descriptivas de variables escalares de acuerdo con el diagnóstico n= 706

	n	Media	DE
Días de hospitalización			
No COVID	240	8.83	9.733
COVID	466	10.63	11.495
Edad (años)			
No COVID	240	45.98	17.817
COVID	466	59.91	12.436
SO2 (%)			
No COVID	148	95.14	3.183
COVID	405	87.46	9.295
Temperatura (°C)			
No COVID	111	36.838	0.8539
COVID	292	37.371	1.048
Frecuencia Cardíaca (lpm)			
No COVID	132	84.21	19.983
COVID	368	83.91	17.503
SOFA (puntos)			
No COVID	158	7.01	3.408
COVID	405	3.39	3.208
Dímero D (µg/ml)			
No COVID	23	2.1848	2.26642
COVID	189	1.6302	1.80885
PaO2/FiO2			
No COVID	143	215.25	93.226
COVID	141	120.25	54.037
Urea (mg/dl)			
No COVID	121	77.07	68.024
COVID	151	89.71	59.277
BUN (mg/dl)			
No COVID	119	34.98	31.062
COVID	150	42.07	28.83
Creatinina (mg/dl)			
No COVID	152	1.393	1.22345
COVID	154	2.0142	2.48714
Proteína C Reactiva (mg/l)			
No COVID	23	186.575	136.9916
COVID	113	127.01	81.958
PEEP (cmH2O)			
No COVID	130	6.39	2.208
COVID	137	8.16	2.33
MAP (cmH2O)			
No COVID	119	11.234	3.3907
COVID	137	14.896	3.4564



Los segmentos de diagonal se generan mediante empates.

Ilustración 1. Curva ROC para Modelo Entrenado



Los segmentos de diagonal se generan mediante empates.

Ilustración 2. Curva ROC para Médico Imagenología

DISCUSIÓN

El mundo de la inteligencia artificial está en constante expansión y ocupa cada vez más espacio en nuestra vida diaria. En el campo de la medicina, el diagnóstico por imagen es un área llena de oportunidades para el desarrollo de la inteligencia artificial, dicha área consiste principalmente en extraer características de las imágenes, interpretar estas características en función del conocimiento recopilado previamente y hacer un diagnóstico. Esta es una serie de pasos que pueden ser simulados de manera eficiente por un sistema informático.

Una de las principales limitantes del uso de redes neuronales convolucionales (CNN) de última generación radica en la necesidad de conjuntos de datos a gran escala para el entrenamiento del modelo y así poder realizar una clasificación y extracción de características precisas.¹⁹

Como se comentó previamente, muchos de los trabajos previos utilizan la misma base de datos e imágenes para entrenar sus modelos y como se evidenció en el trabajo realizado por Magulo, esto favorece la generalización de la clasificación de imágenes de rayos X, debido al hecho de que la red podría aprender características que son específicas del conjunto de datos más no específicas de la enfermedad.²²

Al tratarse de un centro de referencia para pacientes COVID-19 críticamente enfermos, en la ciudad de México, no existe déficit de datos e imágenes para entrenamiento y prueba de un nuevo modelo de inteligencia artificial, por lo que pudimos crear nuestra propia base de datos durante el periodo en que se realizó el trabajo.

En comparación con los estudios previamente publicados, nuestro trabajo utilizó imágenes de tórax, las cuales pudieran considerarse no ideales, imágenes portátiles con proyección anterior posterior, mala calidad de la proyección y gran cantidad de artefactos sin embargo, consideramos que estas mismas características vuelven al modelo más afín a la realidad de unidades de primer y segundo nivel de nuestro país; y pese a dichas características negativas nuestro modelo obtuvo un desempeño similar a lo reportado previamente.

Pese a contar con desempeño diagnóstico similar al reportado en la literatura, obtuvimos resultados de especificidad más bajos; esto pudiera correlacionarse con la gran cantidad

de diagnósticos diferenciales utilizados con la etiqueta “no-COVID”, desde pacientes con pulmones sanos hasta enfermedades infecciosas respiratorias de múltiples causas, edema agudo pulmonar de origen cardiogénico, enfermedades autoinmunes, etcétera. El uso de múltiples diagnósticos en la misma etiqueta, disminuye la capacidad discriminativa del modelo computacional.

Otra característica específica del trabajo realizado fue la inclusión de datos demográficos de los pacientes de quienes se obtuvieron las imágenes, desde la edad y comorbilidades para encontrar poblaciones en mayor riesgo de ser afectados por la enfermedad, hasta características de laboratorio y de la ventilación mecánica que pudieran diferenciarlos de otras causas de distrés respiratorio, no son los objetivos de este estudio sin embargo, el continuar alimentando la base de datos pudiera proporcionar información suficiente para crear un modelo pronóstico de la evolución de pacientes ya afectados por la enfermedad, situación que pudiera ser de mucha utilidad en caso de que los servicios de salud se vean rebasados nuevamente, para la correcta distribución de recursos de las unidades de cuidados intensivos a lo largo del país.

CONCLUSIONES

- 1.- Fuimos capaces de crear un modelo de inteligencia artificial con métodos de aprendizaje máquina utilizando radiografías de tórax en el diagnóstico de pacientes con neumonía por COVID-19 con un rendimiento diagnóstico aceptable con una precisión media (ACCURACY) del 88%.
- 2.- El modelo desarrollado presenta un mejor desempeño como prueba de pesquisa por su sensibilidad alta (99%) y valor predictivo negativo alto (98%).
- 3.- Pese a que ambos observadores (modelo y radiólogo) cuentan con rendimiento diagnóstico similar, el cálculo de concordancia entre ellos resultó negativo.

REFERENCIAS BIBLIOGRAFICAS

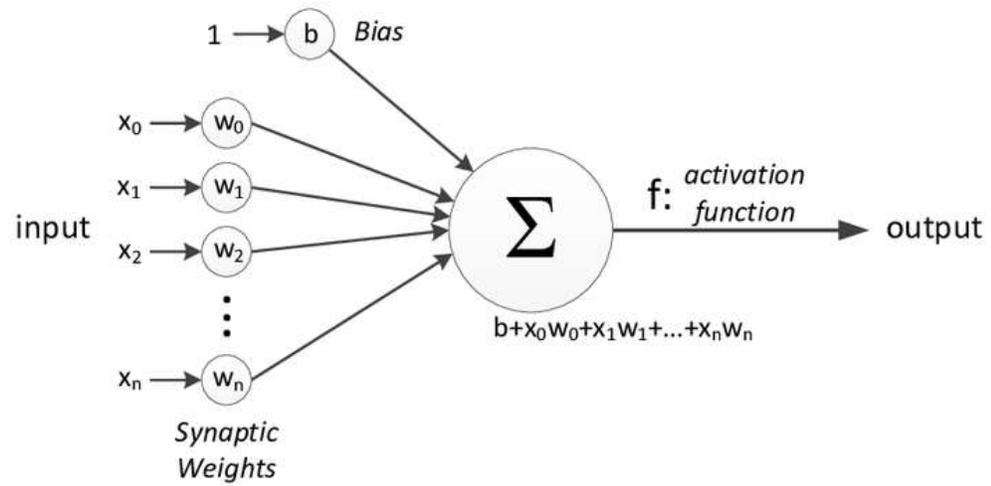
1. Guan W jie, Ni Z yi, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *New England Journal of Medicine*. 2020;382(18):1708-1720. doi:10.1056/nejmoa2002032
2. Normile D. Coronavirus cases have dropped sharply in South Korea. What's the secret to its success? *Science (1979)*. Published online March 17, 2020. doi:10.1126/science.abb7566
3. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*. 2020;20(5):533-534. doi:10.1016/S1473-3099(20)30120-1
4. World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19 - 16 March 2020. WHO. Published March 16, 2020. Accessed March 17, 2020. <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---16-march-2020>
5. Tuite AR, Fisman DN. Reporting, Epidemic Growth, and Reproduction Numbers for the 2019 Novel Coronavirus (2019-nCoV) Epidemic. *Annals of Internal Medicine*. 2020;172(8):567. doi:10.7326/M20-0358
6. Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science (1979)*. 2020;368(6490):489-493. doi:10.1126/science.abb3221
7. Centers for Disease Control and Prevention. Testing Overview. CDC. Published January 21, 2022. Accessed January 21, 2022. <https://www.cdc.gov/coronavirus/2019-ncov/hcp/testing-overview.html#TestingInfection>
8. Tahamtan A, Ardebili A. Real-time RT-PCR in COVID-19 detection: issues affecting the results. *Expert Review of Molecular Diagnostics*. 2020;20(5):453-454. doi:10.1080/14737159.2020.1757437
9. Han SH, Kim KW, Kim S, Youn YC. Artificial Neural Network: Understanding the Basic Concepts without Mathematics. *Dementia and Neurocognitive Disorders*. 2018;17(3):83. doi:10.12779/dnd.2018.17.3.83

10. Liu X, Faes L, Kale AU, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *The Lancet Digital Health*. 2019;1(6):e271-e297. doi:10.1016/S2589-7500(19)30123-2
11. Wang S, Kang B, Ma J, et al. A deep learning algorithm using CT images to screen for Corona virus disease (COVID-19). *European Radiology*. 2021;31(8):6096-6104. doi:10.1007/s00330-021-07715-1
12. Shi H, Han X, Jiang N, et al. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. *The Lancet Infectious Diseases*. 2020;20(4):425-434. doi:10.1016/S1473-3099(20)30086-4
13. Rodriguez-Morales AJ, Cardona-Ospina JA, Gutiérrez-Ocampo E, et al. Clinical, laboratory and imaging features of COVID-19: A systematic review and meta-analysis. *Travel Medicine and Infectious Disease*. 2020;34:101623. doi:10.1016/j.tmaid.2020.101623
14. San Joaquin A. Using Deep Learning to detect Pneumonia caused by NCOV-19 from X-Ray Images. Towards Data Science. Published March 16, 2020. Accessed March 17, 2020. <https://towardsdatascience.com/using-deep-learning-to-detect-ncov-19-from-x-ray-images-1a89701d1acd>
15. Sitaula C, Hossain MB. Attention-based VGG-16 model for COVID-19 chest X-ray image classification. *Applied Intelligence*. 2021;51(5):2850-2863. doi:10.1007/s10489-020-02055-x
16. Thite A. Introduction to VGG16 | What is VGG16? Great Learning Team. Published October 1, 2021. Accessed January 12, 2022. <https://www.mygreatlearning.com/blog/introduction-to-vgg16/>
17. Nishio M, Noguchi S, Matsuo H, Murakami T. Automatic classification between COVID-19 pneumonia, non-COVID-19 pneumonia, and the healthy on chest X-ray image: combination of data augmentation methods. *Scientific Reports*. 2020;10(1):17532. doi:10.1038/s41598-020-74539-2
18. Elaziz MA, Hosny KM, Salah A, Darwish MM, Lu S, Sahlol AT. New machine learning method for image-based diagnosis of COVID-19. *PLOS ONE*. 2020;15(6):e0235187. doi:10.1371/journal.pone.0235187

19. Apostolopoulos ID, Mpesiana TA. Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Physical and Engineering Sciences in Medicine*. 2020;43(2):635-640. doi:10.1007/s13246-020-00865-4
20. Abbas A, Abdelsamea MM, Gaber MM. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *Applied Intelligence*. 2021;51(2):854-864. doi:10.1007/s10489-020-01829-7
21. KC K, Yin Z, Wu M, Wu Z. Evaluation of deep learning-based approaches for COVID-19 classification based on chest X-ray images. *Signal, Image and Video Processing*. 2021;15(5):959-966. doi:10.1007/s11760-020-01820-2
22. Maguolo G, Nanni L. A critic evaluation of methods for COVID-19 automatic detection from X-ray images. *Information Fusion*. 2021;76:1-7. doi:10.1016/j.inffus.2021.04.008
23. Shahinfar S, Meek P, Falzon G. “How many images do I need?” Understanding how sample size per class affects deep learning model performance metrics for balanced designs in autonomous wildlife monitoring. *Ecological Informatics*. 2020;57:101085. doi:10.1016/j.ecoinf.2020.101085
24. Velasco Rodríguez VM, Martínez O VA, Roiz Hernández J. *Muestreo y Tamaño de La Muestra. Una Guía Práctica Para Personal de Salud Que Realiza Investigación*. El Cid Editor; 2003.

1. Red neuronal artificial de una sola capa o perceptrón.

En esta imagen se muestran los diversos elementos de este tipo de redes: Datos de entrada (x), pesos (w), sesgo (b), función de activación y dato de salida (y).



2. Hoja de recolección de datos.



Instituto Mexicano del Seguro Social
UMAE Hospital de Especialidades "Dr. Antonio Fraga Mouret"
Unidad de Cuidados Intensivos



DESARROLLO DE UN MODELO DE INTELIGENCIA ARTIFICIAL CON
MÉTODOS DE APRENDIZAJE MÁQUINA PARA EL DIAGNÓSTICO DE
NEUMONÍA EN PACIENTES CON COVID-19 UTILIZANDO RADIOGRAFÍAS DE TÓRAX.

I. Datos generales.

ID _____ de Imagen: _____

Edad _____ Sexo: H () M ()

Comorbilidades: _____

Diagnóstico _____ Supervivencia: Si () No ()

II. Datos Agregados.

Fecha _____ de imagen: _____

Días _____ desde inicio de síntomas: _____

Días en hospitalización: _____ Ventilación

Mecánica: Si () No ()

Lesión Renal Aguda: Si () No ()

Creatinina sérica: _____ SOFA: _____ PaO2/FiO2: _____

III. Resultado de Análisis.

<i>Observador</i>	COVID	NO COVID
<i>Modelo</i>		
<i>Imagenología</i>		

Comentarios:

