

Aplicación de redes neuronales densas y convolucionales para detección de COVID_19 en imágenes de rayos X
Application of dense and convolutional neural networks for COVID_19 detection in X-ray images

Ronny Stalin Guevara Cruz¹, Claudio Augusto Delrieux²

¹Instituto de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur y Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET), r.s.guevara@hotmail.com, Buenos Aires, Argentina

²Instituto de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur y Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET), cad@uns.edu.ar, Buenos Aires, Argentina

Autor para correspondencia: r.s.guevara@hotmail.com

Fecha de recepción: 2023.05.13

Fecha de aceptación: 2023.06.07

Fecha de publicación: 2023.07.10

RESUMEN

Las redes neuronales convolucionales (CNN) tienen gran potencial en resolver problemas de clasificación con imágenes. La presente investigación tiene como objetivo presentar modelos reducidos que permita identificar casos de neumonía y COVID-19 en imágenes de rayos X de tórax(anterior-posterior), ofreciendo una amplia perspectiva del interés de herramientas que brindan soporte y asistencia médica. La capacidad y tamaño de los modelos fueron reducidos hasta obtener una opción perfecta para ser desplegados localmente en dispositivos con recursos limitados. Los algoritmos propuestos se desarrollaron en Google Colab utilizando el lenguaje de programación Python, aplicando redes neuronales densas y convolucionales a diferentes capas hasta obtener un índice de error bajo, para posterior diagnosticar si el paciente presenta COVID-19. Para ello, se utiliza un conjunto de 603 imágenes de alta resolución de bases de datos públicos (ver en [https://www.cell.com/cell/fulltext/S0092-8674\(18\)30154-5](https://www.cell.com/cell/fulltext/S0092-8674(18)30154-5) y <https://github.com/ieee8023/covid-chestxray-dataset>), divididas en 403 imágenes para entrenamiento, 200 imágenes para prueba y 12 imágenes para validación. La herramienta diseñada con una red neuronal convolucional de 13 capas propone la integración de aprendizaje de máquina (Machine Learning) como soporte en el proceso de diagnóstico médico, con una precisión del 94.73% puede convertirse en una herramienta que brinda mayor velocidad a la hora de dar un diagnóstico.

Palabras clave: COVID-19; neumonía; aprendizaje de máquina; inteligencia artificial; redes neuronales convolucionales.

ABSTRACT

Convolutional neural networks (CNNs) have great potential in solving classification problems with images. The present research aims to present reduced models that allow identifying cases of pneumonia and COVID-19 in chest X-ray images (anterior-posterior), offering a broad perspective of the interest of tools that provide medical support and assistance. The capacity and size of the models were reduced until obtaining a perfect option to be deployed locally in devices with limited resources. The proposed algorithms were developed in Google Colab using the Python programming language, applying dense and convolutional neural networks to different layers until obtaining a low error rate, to later diagnose if the patient has COVID-19. To do this, a set of 603 high-resolution images from public databases (see in [https://www.cell.com/cell/fulltext/S0092-8674\(18\)30154-5](https://www.cell.com/cell/fulltext/S0092-8674(18)30154-5) and <https://github.com/ieee8023/covid-chestxray-dataset>) is used, divided into 403 images for training, 200 images for testing and 12 images for validation. The tool designed with a convolutional neural network of 13 layers proposes the integration of machine learning (Machine Learning) as a support in the medical diagnosis process, with an accuracy of 94.73% can become a tool that provides greater speed when giving a diagnosis.

Key words: COVID-19; pneumonia; machine learning; artificial intelligence; convolutional neural networks

INTRODUCCIÓN

El COVID-19 es una enfermedad infecciosa causada por el síndrome respiratorio agudo severo Coronavirus 2 (SARS-CoV-2), su forma es redonda u ovalada y a menudo polimórfica, con un diámetro de 60 a 140 nm, una enfermedad desconocida que fue descubierta por primera vez en la ciudad de Wuhan, provincia de Hubei, China a finales de 2019 (Pérez Abreu, Gómez Tejeda, & Dieguez Gauch, 2022). Presenta síntomas similares a los de la gripe, entre los que se incluyen fiebre, tos, disnea, mialgia y fatiga, además, causa daños intersticiales pulmonares y, posteriormente, cambios parenquimatosos; en casos más graves se caracteriza por producir neumonía, por esta razón, inicialmente muchos casos de COVID-19 fueron diagnosticados como neumonía (Ye, Zhang, Wang, Huang, & Song, 2020).

La rápida expansión de la enfermedad hizo que la Organización Mundial de la Salud, el 30 de enero de 2020, la declara una emergencia sanitaria de preocupación internacional, basándose en el impacto que el virus podría tener en países subdesarrollados con menos infraestructuras sanitarias y elevó el riesgo de COVID-19 a un nivel muy alto en todo el mundo el 28 de febrero de 2020 (OMS, 2020).

A pesar de los constantes esfuerzos tanto gubernamentales como sociales, al presente se reportan 765 903 278 casos confirmados, 6 927 378 muertes a causa del COVID-19 (OMS, 2023).

En la actualidad, el desarrollo tecnológico permite crear aplicativos capaces de diagnosticar enfermedades rápida y verazmente. Una tecnología de gran utilidad en el desarrollo de sistemas inteligentes son las redes neuronales, puntualmente las redes perceptrón multicapa y redes neuronales convolucionales (MLP & CNNs, por sus siglas en ingles). La MLP es un algoritmo de aprendizaje supervisado, donde una gran cantidad de perceptrones se agrupan en capas y solo existen conexiones directas entre ellos, esto proporciona al algoritmo potentes ventajas como el mapeo no lineal y la tolerancia al ruido (Gómez Macedo, 2022). Mientras que, las CNNs son redes de neuronas que han demostrado su eficiencia en tareas de segmentación y clasificación de imágenes, su función es imitar las neuronas de la corteza visual primaria del cerebro humano (Losada Gutiérrez & Fuentes Jiménez, 2019).

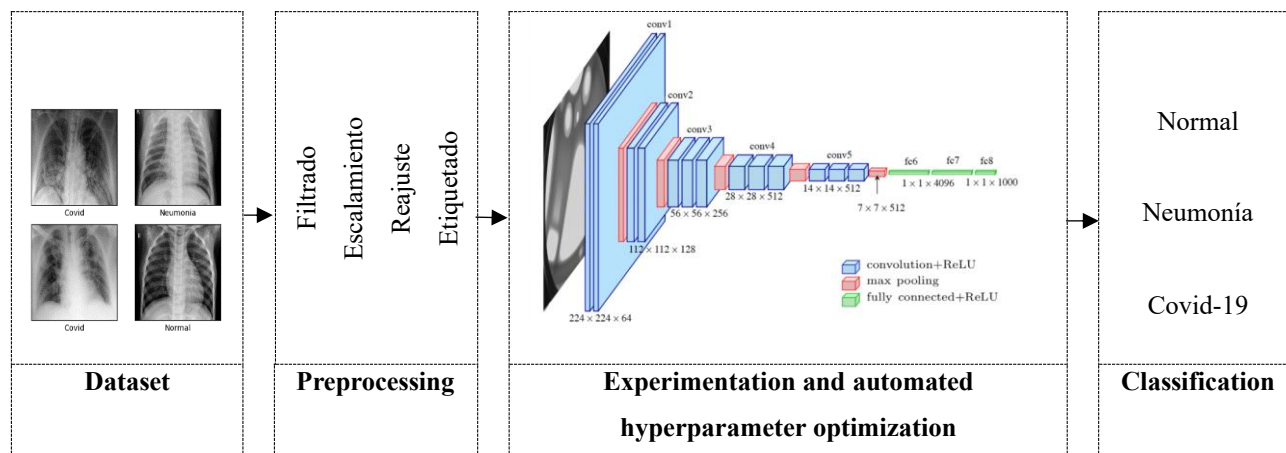
Las arquitecturas de los modelos CNNs y MLPs son capaces de aprender a detectar patrones simples o complejos, en tareas de clasificación de imágenes están compuestas por dos secciones (Lopez Betancur, Bosco Duran, Guerrero Mendez, Zambrano Rodríguez, & Saucedo Anaya, 2021). Específicamente, las CNNs han demostrado ser eficientes en clasificación y detección de afecciones médicas como cáncer de piel (Combalia & Vilaplana, 2019), artritis reumatoide (Serrano Muñoz, Viera López, & Betancourt Hernández, 2018), retinopatía diabética (Catalán, De la Cruz Gámez, Montero Valverde, & Hernández) y en otros casos, de personas en imágenes de profundidad (Losada Gutiérrez & Fuentes Jiménez, 2019). Diversas investigaciones han incursionado en la clasificación de imágenes radiológicas de pecho con COVID-19 (Gómez Macedo, 2022), (Cabanilla del Estal & Martín Martín, 2021) & (Belman Lopez, 2022), sin embargo, carecen de interacción con el usuario generando nuevas posibilidades y desafíos.

En base al análisis realizado, se propone el desarrollo de un aplicativo basada en técnicas Deep Learning aplicando redes neuronales convolucionales con arquitectura ResNet-18 y MLPs-12 para detección de COVID-19, con el objetivo de convertirse en una herramienta que brinde soporte en el diagnóstico médico.

MATERIALES Y MÉTODOS

En la Fig. 1. se muestra la arquitectura del modelo para detección de COVID-19 en imágenes de rayos X, compuesto de una entrada, dos módulos y una salida; el módulo de pre-procesamiento de las imágenes, es el encargo de preparar cada imagen y encapsularla en una tupla para enviarlo al segundo módulo denominado experimentación, encargado de recibir las imágenes, dividirlos en datos para entrenamiento y prueba, ajustar el número de capas y épocas hasta obtener un modelo óptimo tanto en características computacionales como en resultados de predicción.

Fig. 1. Arquitectura del modelo



Diseño del módulo de preprocesamiento

El módulo de pre-procesamiento se encarga de preparar cada imagen y enviarlos al clasificador. Como se muestra en la Fig. 1., su diseño está compuesto de cuatro etapas.

La función `image_processing` encapsula las tres primeras etapas del preprocesamiento de las imágenes. La selección del modelo de color a trabajar con la imagen es fundamental para evitar comportamientos no deseados o retraso en la lectura de la data set. Para tal fin, se elige utilizar la librería de visión artificial y código abierto OpenCV, el comando `cv2.cvtColor` para transformar del espacio RGB a escala de grises, debido a la irregularidad en el tamaño de las imágenes se aplica `cv2.resize` para reajustar la resolución de las imágenes a $600 * 600$ píxeles, `astype` de la librería pandas para cambiar el tipo de dato de la imagen y normalizar en el rango de 0 a 255, finalmente, retornar la información en una tupla.

La etapa de etiquetado crea arreglos de 3 etiquetas, permite al modelo estadístico asociar patrones e imágenes con la etiqueta correcta. Estas corresponden 0 para casos normales, 1 para casos con neumonía y 2 para casos con COVID-19.

Fig. 2. Imagen de rayos X aleatoria resultante del pre-procesamiento



Neumonía

Diseño del módulo clasificación

El módulo de clasificación recibe y entrena al clasificador sirviendo de interfaz con el usuario, a fin de, visualizar los resultados obtenidos. Como se muestra en la Fig. 1., su diseño está compuesto de 2 etapas y un submódulo de retroalimentación.

En este trabajo se propone el uso y adaptación de una MLP con 12 capas y una ResNet a 18 capas. La etapa MLP (multi layer perceptron) utiliza una red neuronal unidireccional constituida por tres o más capas: una capa de entrada, una capa de salida y capas intermedias o capas ocultas (Quiñones Huatangari, Ochoa Toledo, & Gamarra Torres, 2020). Para el desarrollo de la investigación se consideran inicialmente como hiperparámetros:

Tabla 1. Hiperparámetros para la red MLP de 4 capas

Parámetros / Hiperparámetros	Valor	Activación
Capas de entrada	1	
Capas densas	4	Relu
Capas de salida	1	Softmax
Épocas	30	
Optimizador	ADAM	
Tasa de aprendizaje	0.001	
Batch size	13	

Con la finalidad de mejorar el clasificador se aumenta el número de capas Dense con arquitectura encoders y las épocas de entrenamiento de la red, obteniendo:

Tabla 2. Hiperparámetros para la red MLP de 12 capas

Parámetros / Hiperparámetros	Valor	Activación
Capas de entrada	1	
Capas densas	12	Relu
Capas de salida	1	Softmax
Épocas	30	
Optimizador	ADAM	
Tasa de aprendizaje	0.001	
Batch size	13	

La etapa CNN (convolutional neural networks) utiliza la operación de convolución como base para el procesamiento de datos. Dentro de esta red, existen matrices llamadas filtros y cada una de ellas detecta, a su vez, distintos tipos de características dentro de los datos que se quieren procesar. En imágenes estos filtros son capaces de detectar bordes, moviéndose según un parámetro que mide la longitud del salto (stride) (Bueno, 2019). Para el desarrollo de la investigación se consideran como hiperparámetros:

Tabla 3. Hiperparámetros para la red CCN de 18 capas

Parámetros / Hiperparámetros	Valor	Activación
Capas de entrada	1	
Capas de convolución	8	Relu
Capas densas	8	Relu
Capas de activación	1	Dropout
Capas de salida	1	Softmax
Épocas	75	
Optimizador	ADAM	
Tasa de aprendizaje	0.001	
Batch size	13	

El submódulo de retroalimentación permite ajustar los clasificadores hasta obtener un modelo óptimo tanto en características computacionales como en resultados de predicción, mediante el análisis de las métricas de desempeño y la magnitud de pérdida en razón del número de épocas de entrenamiento.

RESULTADOS Y DISCUSIÓN

Ajuste del clasificador MLP

Para alcanzar el objetivo funcional del aplicativo se realizaron pruebas en el ajuste de capas y épocas. El procedimiento inicia con 4 capas densas y 30 épocas de entrenamiento para verificar las métricas de desempeño, mediante el análisis de la matriz de confusión (Fig. 3.), las métricas de evaluación se muestran en la Tabla 4.

Tabla 4. Métricas de evaluación, MLP 4 capas

	precision	recall	f1-score	support
0	0.9000	0.9114	0.9057	79
1	0.8537	0.7955	0.8667	44
2	0.9367	0.9610	0.9487	77
accuracy			0.9050	200
macro avg	0.8968	0.8893	0.8926	200
weighted avg	0.9039	0.9050	0.9042	200

Fig. 3. Matriz de confusión, MLP 4 capas

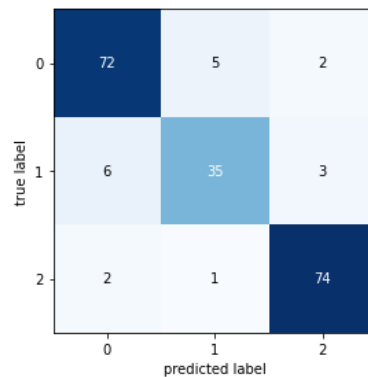
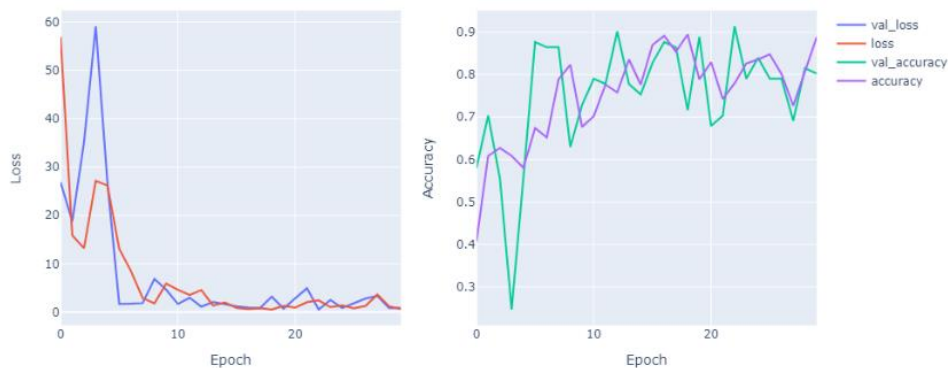


Fig. 4. Perdida y exactitud durante la trama de entrenamiento red MPLs



Para este modelo denso sin capas de convolución el test con 200 imágenes obtuvo los siguientes resultados:

- 72 casos donde el paciente no tiene neumonía ni COVID y el algoritmo predijo que es la imagen de un tórax normal, True Positive (TP)
- 35 casos donde el paciente tenía neumonía y el algoritmo predijo que es la imagen de un tórax con neumonía, True Positive (TP)
- 74 casos donde el paciente tenía COVID-19 y el algoritmo predijo que es la imagen de un tórax con COVID-19, True Positive (TP)
- 19 casos donde el algoritmo clasifico erróneamente.

El clasificador se ajusta a 13 capas densas siguiendo el modelo de arquitectura encoders y 50 épocas de entrenamiento, la Tabla 5 muestra los resultados.

Tabla 5. Métricas de evaluación, MLP 12 capas

	precision	recall	f1-score	support
0	0.9333	0.8861	0.9091	79
1	0.8125	0.8864	0.8478	44
2	0.9610	0.9610	0.9610	77
accuracy			0.9150	200
macro avg	0.9023	0.9112	0.9060	200
weighted avg	0.9174	0.9150	0.9156	200

Fig. 5. Matriz de confusión, MLP 12 capas

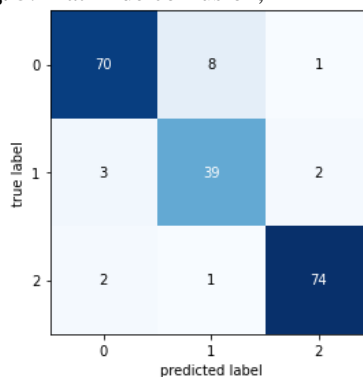
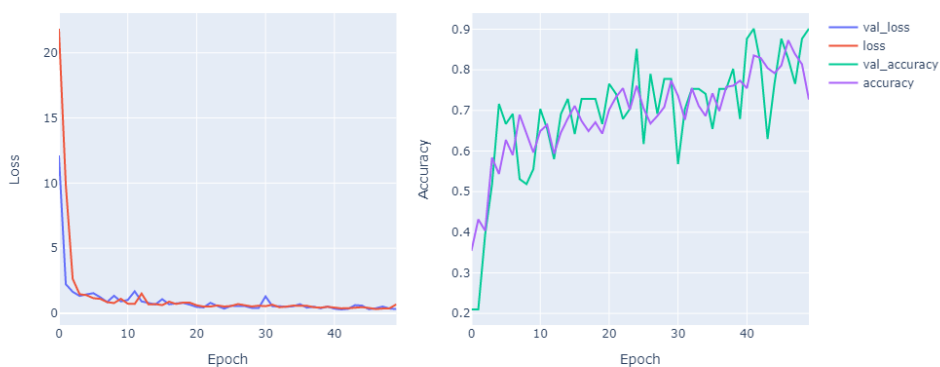


Fig. 6. Perdida y exactitud durante la trama de entrenamiento red MPLs



Para este modelo denso sin capas de convolución el test con 200 imágenes obtuvo los siguientes resultados:

- 70 casos donde el paciente no tiene neumonía ni COVID y el algoritmo predijo que es la imagen de un tórax normal, True Positive (TP)
- 39 casos donde el paciente tenía neumonía y el algoritmo predijo que es la imagen de un tórax con neumonía, True Positive (TP)
- 74 casos donde el paciente tenía COVID-19 y el algoritmo predijo que es la imagen de un tórax con COVID-19, True Positive (TP)
- 17 casos donde el algoritmo clasifico erróneamente.

Ajuste del clasificador CNN

La selección del número de épocas se realizó en base al comportamiento de aprendizaje del modelo, a fin de evitar que la CNN alcance overfitting, y que aprenda del ruido (Lopez Betancur, Bosco Duran, Guerrero Mendez, Zambrano Rodríguez, & Saucedo Anaya, 2021). Para que el sistema sea más robusto se incrementa la cantidad de datos utilizando las siguientes transformaciones: RandomResizedCrop donde la imagen es recortada aleatoriamente, y renderizada de acuerdo con los tamaños requeridos por el clasificador, rotation_range rota aleatoriamente la imagen dentro de un rango determinado (Lopez Betancur, Bosco Duran, Guerrero Mendez, Zambrano Rodríguez, & Saucedo Anaya, 2021). En esta investigación se seleccionó un rango de 0 a 30 grados. Zoom_range renderiza una nueva imagen con un rango de acercamiento, rango de 0.5 a 1.5. Width_shift_range y height_shift_range desplaza la imagen de manera horizontal y vertical respectivamente.

Tabla 6. Métricas de evaluación, CNN 18 capas

	precision	recall	f1-score	support
0	0.8721	0.9494	0.9091	79
1	0.9459	0.7955	0.8642	44
2	0.9481	0.9481	0.9481	77
accuracy			0.9473	200
macro avg	0.9220	0.8976	0.9171	200
weighted avg	0.9176	0.9150	0.9242	200

Fig. 7. Matriz de confusión, CNN 18 capas

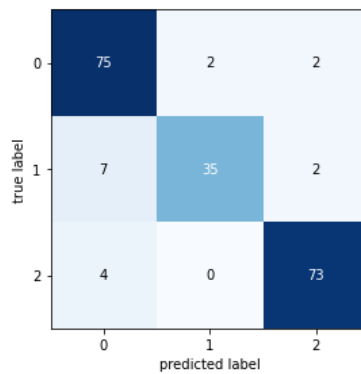
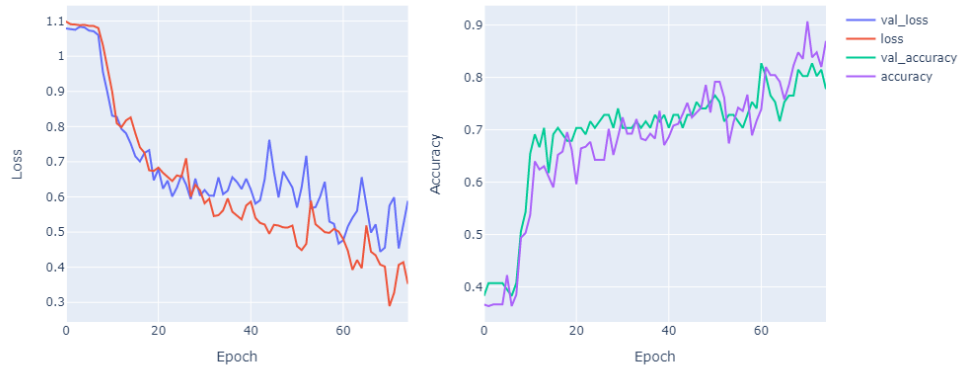


Fig. 8. Perdida y exactitud durante la trama de entrenamiento red MPLs



Para el modelo denso con capas de convolución el test con 200 imágenes obtuvo el siguiente resultado:

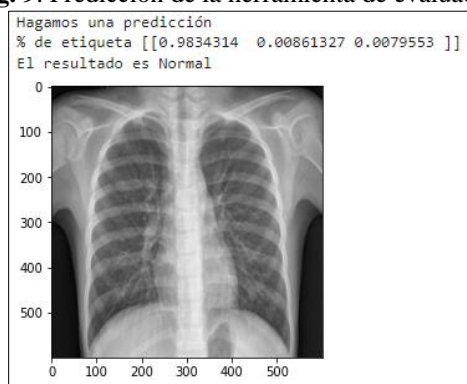
- 75 casos donde el paciente no tiene neumonía ni COVID y el algoritmo predijo que es la imagen de un tórax normal, True Positive (TP)
- 35 casos donde el paciente tenía neumonía y el algoritmo predijo que es la imagen de un tórax con neumonía, True Positive (TP)

- 73 casos donde el paciente tenía COVID-19 y el algoritmo predijo que es la imagen de un tórax con COVID-19, True Positive (TP)
- 17 casos donde el algoritmo clasifico erróneamente.

Validación

La evaluación del sistema fue realizada mediante pruebas de funcionamiento. Se desarrollo un aplicativo que permite al usuario cargar y seleccionar la imagen de rayos X que requiere evaluar. En cada interacción se imprime el resultado y correspondiente etiqueta.

Fig. 9. Predicción de la herramienta de evaluación



DISCUSIÓN

En la fase de ajuste del clasificador MLP se observó que al incrementar el número de capas densas ocultas esta alcanzaba overfitting y una notable reducción de precisión respecto a un clasificador de 12 capas, ambas entrenadas con 30 épocas. En este sentido, para esta investigación se propone un modelo de clasificación MPLs-12 basado en la arquitectura de autoencoder. Al analizar los datos mostrados en la sección 3.1 se puede observar una exactitud de 91,50% esta puede ser aplicable para clasificación de imágenes, sin embargo, no se considera suficiente para clasificaciones de imágenes médicas. El modelo presenta una importante desventaja, si la imagen evaluada sufre alguna alteración en su tamaño o forma, el clasificador tiende a fallar.

Para solventar el inconveniente, se propone modificar el modelo ResNet-50 hasta obtener un modelo optimizado a dispositivos de capacidades limitadas, los resultados expuestos en la sección 3.2 proponen un modelo de 18 capas con una exactitud de 94.73%.

Después de analizar los datos obtenidos, se determinó que la ResNet-18 cuenta con una mejor exactitud y precisión al clasificar escenarios de tres salidas sin ocasionar saturación a dispositivos de recursos limitados. La sección 3.3 denota lo anteriormente expuesto.

CONCLUSIONES

En esta investigación se presenta la comparación entre nuevos modelos de redes neuronales densas y convolucionales con arquitecturas DenseNet-121 y ResNet-50 como clasificadores de COVID-19, con el objetivo de determinar la más confiable para este diagnóstico y convertirla en una herramienta que brinde asistencia y soporte médico.

Los modelos propuestos fueron desarrollados para proporcionar diagnósticos precisos en más clases de salida que estudios anteriores solo cubren escenarios binarios. Para mejorar la precisión en los resultados y prevenir el sobreajuste a los modelos se agregan técnicas de regularización como dropout y un correcto ajuste de hiperparámetros. Las arquitecturas pre-entrenadas extraídas de las librerías Keras y Tensorflow de código abierto, utilizan imágenes de rayos X frontales de tórax para el entrenamiento, su tamaño fue reducido con la finalidad de implementarse en dispositivos con capacidades limitadas obteniendo MPLs-12 y ResNet-18.

La red neuronal convolucional con arquitectura ResNet-18 entrenada con 75 épocas es capaz de clasificar el COVID-19 con una precisión del 94.73%, lo que nos indica un excelente desempeño de predicción y alto grado de clasificación. Respecto al tiempo total de entrenamiento, la MPLs-12 se ejecuta en menor tiempo, aunque su resultado de precisión de clasificación alcanza un 91,50%, si la imagen a evaluar sufre una alteración en sus ejes, el modelo no predice de manera asertiva.

La presente investigación puede servir de base para el desarrollo informático de aplicaciones de apoyo para toma de decisiones y diagnósticos. Además de la clasificación se propone emplear técnicas de segmentación de imágenes para detección exacta del daño en el pulmón en colaboración directa de los centros médicos y sus bases de datos para el entrenamiento continuo y mejoramiento de resultados.

Debemos recordar que el COVID-19 inicio como una pandemia donde el conocimiento fue limitado y ahora que contamos con datos debemos mejorar el diagnostico de quienes diariamente arriesgaron sus vidas.

REFERENCIAS

- Belman Lopez, C. E. (2022). *Detection of COVID-19 and Other Pneumonia Cases using Convolutional Neural Networks and X-ray Images* (Vol. 42). Colombia: Ingeniería e Investigación. doi:<https://doi.org/10.15446/ing.investig.v42n1.90289>
- Bueno, F. (2019). *Redes Neuronales: Entrenamiento y Comportamiento*. Madrid: Universidad Complutense Madrid.
- Cabanilla del Estal, T., & Martín Martín, Q. (2021). *Aplicación de las redes neuronales artificiales al COVID-19*. España: Universidad de Salamanca.
- Catalán, E., De la Cruz Gámez, E., Montero Valverde, J., & Hernández, R. (s.f.). *Detección automática de retinopatía diabética aplicando visión artificial y redes neuronales convolucionales*.
- Combalia, M., & Vilaplana, V. (2019). *Clasificación de imágenes dermatoscópicas utilizando Redes Neuronales Convolucionales e información de metadatos*. Barcelona: Universidad Politécnica de Catalunya.
- Gómez Macedo, M. (2022). *Detección de covid-19 y neumonía en imágenes rayos x de pulmones utilizando redes neuronales convolucionales*. México: Universidad Nacional Autónoma de México.
- Lopez Betancur, D., Bosco Duran, R., Guerrero Mendez, C., Zambrano Rodríguez, R., & Saucedo Anaya, T. (2021). Comparación de arquitecturas de redes neuronales convolucionales para el diagnóstico de COVID-19. En *Computacion y Sistemas* (Vol. 25, págs. 601–615). Mexico: Instituto Politécnico Nacional/Centro de Investigacion en Computación. doi:10.13053/CyS-25-3-3453
- Losada Gutiérrez, C., & Fuentes Jiménez, D. (2019). *Detección de personas en imágenes de profundidad mediante redes neuronales convolucionales*. Madrid: Universidad de Alcalá.
- OMS. (2020). *Coronavirus disease 2019 (COVID-19). Situation Report – 39*. Geneva: World Health Organization. Recuperado el 30 de Agosto de 2022, de

https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200228-sitrep-39-covid-19.pdf?sfvrsn=5bbf3e7d_2

OMS. (2023). *Coronavirus disease (COVID-19) pandemic*. Recuperado el 2022 de Septiembre de 02, de <https://www.who.int/emergencias/diseases/novel-coronavirus-2019>

Pérez Abreu, M. R., Gómez Tejeda, J. J., & Dieguez Gauch, R. A. (2022). *Características clínico-epidemiológicas de la COVID-19* (Vol. 19). La Habana: Revista Habanera de Ciencias Médicas. Obtenido de <http://www.revhabanera.sld.cu/index.php/rhab/article/view/3254/2505>

Quiñones Huatangari, L., Ochoa Toledo, L., & Gamarra Torres, O. (2020). *Red neuronal artificial para estimar un índice de calidad de agua*. Quito: Enfoque. doi:<https://doi.org/10.29019/enfoque.v11n2.633>

Serrano Muñoz, A., Viera López, G., & Betancourt Hernández, M. (2018). *Diagnostico automático de artritis reumatoide en radiografías de manos utilizando redes neuronales convolucionales*. (Vol. 35). Cuba: Revista Cubana de Física.

Ye, Z., Zhang, Y., Wang, Y., Huang, Z., & Song, B. (2020). Chest CT manifestations of new coronavirus disease 2019 (COVID-19): a pictorial review. En *European Radiology* (Vol. 30, págs. 4381-4389). doi:10.1007/s00330-020-06801-0