

Sistema Auxiliar para el Diagnóstico de COVID-19 Mediante Análisis de Imágenes de CR Torácica Basado en Deep Learning

I. Perrilliat-García^{1§}, M. A. Gámez-Guerrero^{1§},
S. L. Rocha-Nava^{1,2}, J. I. Hernández-Oropeza^{1*}

¹ Universidad La Salle, Ciudad de México, México.

² Instituto Nacional de Cancerología (INCan), Ciudad de México, México.

§ Estos autores contribuyeron de manera equitativa en este trabajo.

* jose.hernandez@lasalle.mx

Resumen

La COVID-19 es la enfermedad causada por el coronavirus, SARS-CoV-2, descubierto en Wuhan, China, a finales de 2019, el cual originó una pandemia que afecta hoy en día a todo el mundo. En este sentido, los sistemas de salud se encuentran bajo constante presión, por lo que la necesidad de establecer nuevos protocolos y herramientas más eficientes para acortar los tiempos de diagnóstico y tratamiento resulta evidente. La COVID-19 puede ser diagnosticada por pruebas moleculares o por imágenes médicas. En este sentido, el diagnóstico por Radiografía Computarizada (CR, por sus siglas en inglés) cuenta con algunas ventajas para ser tomado en cuenta para la detección oportuna del COVID-19. En este trabajo, se presentan los resultados de la aplicación de redes neuronales convolucionales para el desarrollo de un sistema capaz de clasificar imágenes de CR torácicas de pacientes sanos, con neumonía y COVID-19 con una precisión de clasificación global mayor al 95 % a partir del uso de una base de datos de 1800 imágenes.

Palabras clave: Sistema auxiliar, Herramienta computacional, COVID-19, CR torácica, Deep Learning.

1. Introducción

Los coronavirus pertenecen a una vasta familia de virus que afectan a la salud de los humanos. La Organización Mundial de la Salud (OMS) reporta la existencia de varios tipos de coronavirus que pueden causar infecciones respiratorias. En el caso específico del SARS-CoV-2, descubierto en Wuhan, China en diciembre del 2019, causa la enfermedad denominada COVID-19 [1]. Se sabe que este coronavirus tiene una gran facilidad de propagación, lo cual ha provocado un alto índice de contagios y una pandemia que ha tenido impactos negativos no solo en la salud, sino en lo económico, social y emocional de la población mundial.

Actualmente, el diagnóstico de COVID-19 se puede realizar mediante los siguientes procedimientos:

- Reacción en Cadena de la Polimerasa con Transcriptasa Inversa (por sus siglas en inglés, RT-PCR), la cual es un método en el que se amplifican regiones específicas del genoma del coronavirus SARS-CoV-2 [2]. La desventaja de esta prueba es que puede tardar en ser entregado el resultado hasta 2 días y requiere de instalaciones con un ambiente de prueba altamente controlado.

- Pruebas serológicas, si bien son útiles para confirmar la respuesta inmunológica a un patógeno viral perteneciente al coronavirus, no son muy recomendables debido a que los rangos de sensibilidad y especificidad reportados, 88.66 % y 90.63 %, respectivamente, pueden arrojar un porcentaje significativo de falsos negativos y positivos [3].
- Radiografía Computarizada (CR, por sus siglas en inglés), permite el diagnóstico utilizando imágenes obtenidas por la aplicación de rayos X a través del cuerpo del paciente.
- Tomografía Computarizada (CT, por sus siglas en inglés), el diagnóstico se realiza por medio de imágenes obtenidas por cortes de las estructuras del paciente. Este método requiere un diseño arquitectónico y de instalaciones especiales capaces de soportar el peso del equipo y sus accesorios, además de un blindaje de plomo especial, así como contar con las características de suministro eléctricas adecuadas y un sistema adicional de enfriamiento entre otros [4], las cuales serían difíciles de incluir en áreas reacondicionadas para tratar pacientes con COVID-19.

Está reportado que la detección del COVID-19 por medio de diagnósticos radiológicos (CT) presenta mayor sensibilidad, a comparación de otras técnicas como la RT-PCR [5, 6]. Sin embargo, el diagnóstico por CR cuenta con algunas ventajas para ser tomado en cuenta para la detección oportuna del COVID-19. En primer lugar, este estudio se realiza en un menor tiempo, por lo tanto, hay un ahorro de tiempo crítico para el diagnóstico, lo que representa un menor costo. Además, es más factible contar con un equipo de rayos X que con uno de CT dentro de las unidades de atención médica, así mismo no es despreciable mencionar que la tasa de radiación es menor y más sencillo de utilizar que un diagnóstico radiológico utilizando CT [7].

Por otra parte, el Deep Learning se puede describir como una técnica de Inteligencia Artificial (AI, por sus siglas en inglés), del subcampo conocido como Machine Learning, mediante el cual las computadoras son capaces de emular el aprendizaje humano, con la finalidad de automatizar sistemas predictivos y representar datos a través de un algoritmo basado en capas. A partir de la cantidad de capas que forman un modelo se determina la profundidad de este. Mayormente estos modelos de datos se basan en las llamadas redes neuronales artificiales que se encargan del procesamiento de información a través de una estructura de varias capas conjuntas [8]. Estos sistemas están principalmente formados por tres partes o capas: capa de entrada (*Input Layer*), capa oculta (*Hidden Layer*) y capa de salida (*Output Layer*). Estos algoritmos se pueden clasificar de manera general en cuatro tipos: de aprendizaje supervisado, no supervisado, por refuerzo y auto-supervisado [8]. En este proyecto, un sistema auxiliar para el diagnóstico de COVID-19 es propuesto. El sistema está basado en un modelo de aprendizaje profundo supervisado, ya que se cuenta con los datos de entrada (imágenes de CR torácicas) y sus anotaciones o etiquetas proporcionadas por los médicos; es decir, se cuenta con el diagnóstico obtenido a partir de cada imagen.

Cabe mencionar que la aplicación de estos modelos es utilizada especialmente para resolver problemas de clasificación y regresión, pero también existen los modelos de red neuronal convolucional que son especialmente potentes para resolver tareas de visión por computadora [8], lo cual es útil para cumplir con el objetivo principal de este trabajo, que fue determinar si la imagen cargada al sistema mostraba rasgos característicos asociados con imágenes de CR torácica de pacientes diagnosticados con COVID-19.

2. Metodología

La metodología empleada para el desarrollo del sistema auxiliar propuesto en este trabajo consistió en cuatro etapas, las cuales se describen a continuación:

A. *Recopilación.* Las imágenes utilizadas en este proyecto fueron tomadas de *Kaggle*, una comunidad de científicos de datos y profesionales del aprendizaje automático [9, 10] de la cual se sustrajeron 1800 imágenes, de las cuales 600 fueron de pulmones de pacientes diagnosticados con COVID-19, 600 imágenes de pulmones de pacientes diagnosticados con neumonía y 600 imágenes de pulmones sanos. Para los fines de este proyecto se asumieron como verdaderos los datos proporcionados por los autores de cada base de datos, ya que al ser de acceso público cada uno cuenta con su propia metodología de validación. En la Fig. 1, se muestra un ejemplo para cada clase o categoría de las imágenes utilizadas.

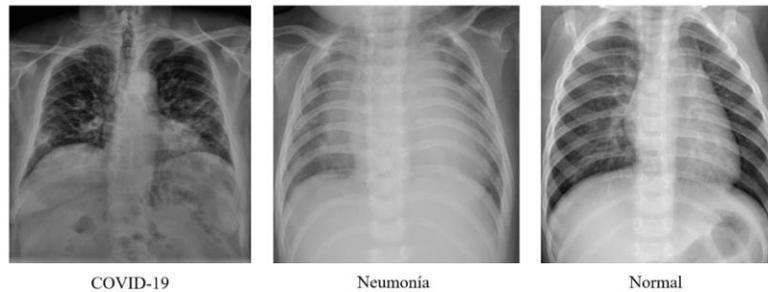


Fig. 1. Comparación de las 3 clases empleadas en el sistema.

B. *Estructuración y normalización.* Las imágenes originales de una resolución promedio de 2200 x 2200 píxeles en distintos formatos (JPG y PNG), fueron normalizadas a una resolución de 224 x 224 píxeles y a formato JPG, posteriormente fueron repartidas en la proporción 80/20 (1350 para *Train* y 450 para *Validation*) para el modelo 1, y 75/15/15 (1260 para *Train*, 270 para *Validation* y 270 para *Test*) para el modelo 2. Se formaron dos árboles de carpeta o *DataSets* como se muestra en la Fig. 2.

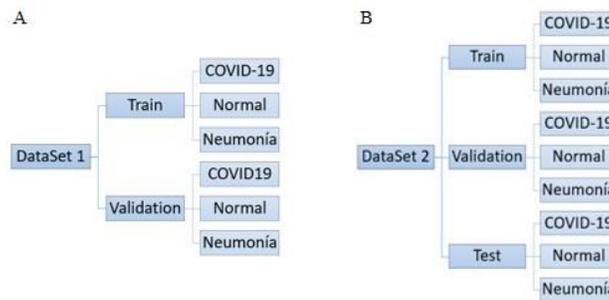


Fig. 2. A) Árbol de carpetas del modelo 1. B) Árbol de carpetas del modelo 2.

C. *Diseño de la arquitectura de red neuronal artificial.* El sistema se realizó basado en el lenguaje de programación *Python*, en el entorno de desarrollo *Google Colaboratory* que permite ejecutar programas con acceso gratuito a Unidades de Procesamiento Gráfico (*GPU*, por sus siglas en inglés). Fueron diseñadas dos arquitecturas de red neuronal artificial específicas para cada árbol de carpetas, capaces de procesar las imágenes de manera convolucional, para esto se exploraron los resultados obtenidos a partir de la combinación de distintos parámetros como lo son el tipo de capa, tamaño de kernel, número de capas, número de neuronas por capa, función de coste, optimizadores, tasa de aprendizaje entre otros, a lo largo de un gran número de ensayos hasta obtener los mejores resultados.

D. Predicción y validación. Una vez obtenidos los modelos de redes neuronales convolucionales, fueron utilizados para realizar predicciones con los datos de la carpeta *Validation* en el caso del modelo 1 y *Test* para el modelo 2. Los resultados fueron comparados contra las clases verdaderas o diagnósticos de cada imagen en una matriz de confusión, la cual permite visualizar la cantidad de Verdaderos Positivos (*TP*)/ Verdaderos Negativos (*TN*), así como los Falsos Positivos (*FP*) / Falsos Negativos (*FN*).

Con los resultados de predicción, fueron calculados los valores de precisión, exactitud, especificidad, y sensibilidad de los sistemas, además del cálculo del valor-F1 o *F1-score*. Las métricas de validación se determinaron acorde a lo reportado por [11, 12, 13], tomando en cuenta los valores de *TP*, *TN*, *FP* y *FN*.

3. Resultados y Discusión

En la Fig. 5A, se muestran los resultados obtenidos al entrenar la red neuronal convolucional propuesta con el *DataSet 1* junto con sus valores de exactitud obtenidos durante el entrenamiento (*train acc*) y validación (*val acc*). Por otra parte, en la Fig. 5B, se presentan los resultados obtenidos al entrenar la red neuronal convolucional propuesta con el *DataSet 2*, como se puede observar, los resultados de exactitud obtenidos durante el entrenamiento (*train acc*), validación (*val acc*) y prueba (*test acc*) fueron superiores, por lo cual se decidió seguir solamente con el modelo 2.

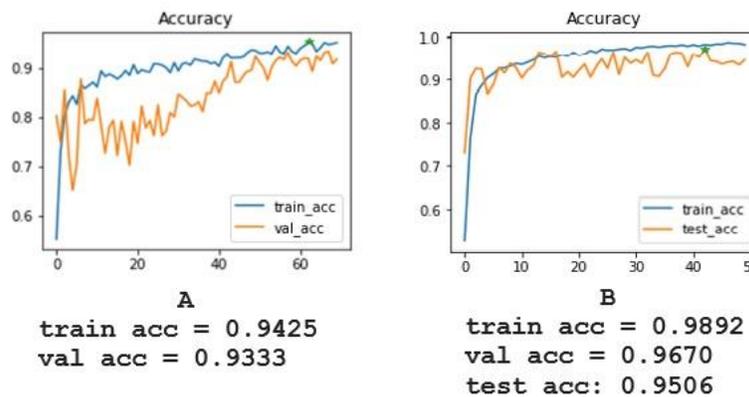


Fig. 5. Curvas de aprendizaje y valores de exactitud o *accuracy* de A) (Entrenamiento y Validación) del modelo 1 y B) (Entrenamiento, Validación y Prueba) del modelo 2.

En la Fig. 6, se muestra la matriz de confusión para la carpeta *Test* con sus respectivas predicciones de las clases "COVID-19", "Neumonía" y "Normal" seguido de los cálculos obtenidos de precisión, exactitud, sensibilidad, especificidad y valor-F1 mostrados en la Tabla 2. Cabe resaltar que esta carpeta fue reservada para validar la precisión mostrada en las curvas de aprendizaje, al ser imágenes totalmente nuevas para la red neuronal convolucional propuesta.

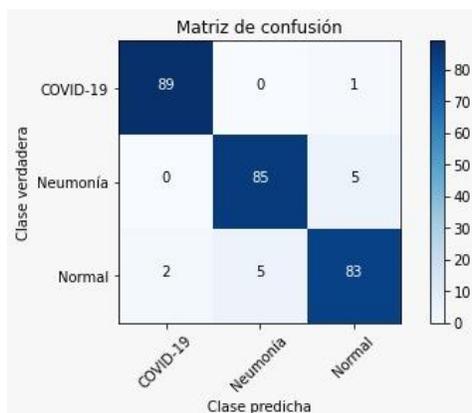


Fig. 6. Matriz de confusión de la red neuronal convolucional para la carpeta Test del DataSet 2.

Como se puede observar en la Tabla 2, el sistema propuesto presenta un porcentaje de precisión mayor al 97 % para la detección de COVID-19, mayor al 94 % para la detección de neumonía y mayor al 93 % para la detección de pulmones normales. Comparando estos valores con lo reportado por [14], se puede decir que ambos modelos presentan niveles similares de sensibilidad con la clase “Normal” y “No COVID-19”, que para el caso del modelo 2 se denominó como “Neumonía”. No obstante, la sensibilidad hacia la clase “COVID-19” del modelo 2 del presente trabajo es mayor en un 7.8 %.

Así mismo, la precisión del modelo 2 es menor con la clase “COVID-19”, pero presenta resultados mayores con las clases “No COVID-19 o Neumonía” y “Normal”, con un 3.14 % y 2.75 %, respectivamente. El sistema realizó la predicción o clasificación del conjunto *Test* en un promedio de 0.3068 segundos, considerando 10 ensayos.

Tabla 2: Métricas de evaluación del modelo de red neuronal convolucional

Clases	Precisión	Exactitud	Sensibilidad	Especificidad	Valor F1
COVID-19	97.802 %	98.889 %	98.889 %	98.889 %	98.343 %
Neumonía	94.444 %	96.296 %	94.444 %	97.222 %	94.444 %
Normal	93.258 %	95.185 %	92.222 %	96.667 %	92.737 %
Promedio	95.168 %	96.790 %	95.185 %	97.593 %	95.175 %

Finalmente, las métricas obtenidas por el modelo 2 muestran un mejor desempeño en los promedios globales de precisión (5.208 % mayor), exactitud (9.77 % mayor), sensibilidad (9.835 % mayor), especificidad (5.413 % mayor) y valor-F1 (7.805 % mayor) que las reportadas por [15], haciéndolo más eficiente al momento de realizar las predicciones de las 3 clases establecidas.

4. Conclusiones

El sistema propuesto fue capaz de diferenciar imágenes de CR torácicas de pacientes con neumonía, COVID-19 y sanos, con un porcentaje mayor al 95 %, lo cual indica que este sistema es capaz de identificar patrones y características específicas de las enfermedades establecidas para este trabajo.

Las curvas de aprendizaje presentan una estabilización a lo largo de las épocas, a pesar de mostrar un ligero sobreajuste (*Overfitting*) durante la validación del modelo 2. Sin embargo, este no representa un problema grave en los resultados de exactitud obtenidos al final del entrenamiento (Entrenamiento (*train acc*) = 0.9892; Validación (*val acc*) = 0.9670) ya que fueron favorables los resultados para predicciones posteriores (Prueba (*test acc*) = 0.9506).

El uso de *Data augmentation* es una estrategia que ayuda a simular una cantidad mayor de datos de los que se tienen para optimizar el entrenamiento de modelos, incluyendo características aleatorias en cada iteración, en este caso se utilizaron giros horizontales y una variación de zoom en las imágenes, mostrando un aumento en los porcentajes de precisión dentro de las curvas de aprendizaje, así como fue empleado y reportado por [16].

Actualmente, este sistema se maneja con el código basado en *Python*, el cual contiene una función que permite normalizar cualquier imagen al formato y tamaño necesarios para realizar la clasificación de esta. De igual forma, el sistema puede ser utilizado en un equipo promedio con una debida instalación y sin conexión a Internet, ya que este recurso no es necesario para realizar la clasificación de las imágenes a escala clínica. Así mismo, dada la escalabilidad del sistema, se buscaría implementar una interfaz gráfica para facilitar su uso por parte del usuario final.

Este sistema auxiliar es una herramienta computacional que podría brindar una solución sencilla pero ágil, en el sentido de que la emergencia sanitaria actual exige soluciones inmediatas, al disminuir considerablemente la carga de trabajo en los hospitales, ahorrando tiempo vital en el diagnóstico de varios pacientes al ser solicitados estudios de imagen para evaluar el desarrollo de la enfermedad, al descartar de manera automática los casos negativos o normales. Así mismo, este sistema podría auxiliar a los médicos no expertos para la interpretación de CR torácicas, especialmente en aquellas unidades adaptadas contra el COVID-19 donde no se tenga la infraestructura adecuada para el uso de un equipo de CT. Cabe resaltar que este sistema podría mejorarse si se utilizaran imágenes médicas proporcionadas por un hospital y un método de validación adicional que comparara la concordancia entre la clasificación arrojada por el sistema y el diagnóstico emitido por un médico radiólogo experto para cada imagen utilizada en la carpeta Test.

Declaración de conflictos de interés

Los autores declaran no tener ningún conflicto de interés para este trabajo.

Agradecimientos

Los autores agradecen a Dr. José Ambrosio Bastián, Dr. Adrián Octavio Ramírez Morales y Mtro. Juan Francisco López Luna, miembros de la Academia de Bioingeniería de la Licenciatura en Ingeniería Biomédica de la Universidad La Salle, por haber revisado y expresado sugerencias para la mejora del presente trabajo.

Asimismo, un especial agradecimiento a la Facultad de Ingeniería y a la Dirección de Investigación y Posgrado de la Universidad La Salle por el financiamiento otorgado para la realización de este proyecto.

Referencias

- [1] A. Cossarizza, S. De Biasi, G. Guaraldi, M. Girardis, and C. Mussini, "SARS-CoV-2, the virus that causes COVID-19: Cytometry and new challenge for global health," *Cytometry Part A*, 97A, pp. 340-343, 2020.
- [2] C. Long, H. Xu, Q. Shen, X. Zhang, B. Fan, et al., "Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT?," *Eur. J. Radiol.*, vol. 126, no. 108961, pp. 1-6, 2020.
- [3] J. M. Meza-Calvache, A. D. Estrada-Rodríguez, C. B. Chabusa-Martínez, and V. A. Velasco-Paucar, "Utilidad de Pruebas de cadena de polimerasa, pruebas rápidas y Tomografías en pacientes con Covid-19," *J. Am. Health*, vol. 3, no. 2, pp. 32-39, 2020.
- [4] Guía Tecnológica No. 6: Tomografía Computarizada, CENETEC, Secretaría de Salud, Ciudad de México, México, 2004, pp. 4-6.
- [5] Y. Fang, H. Zhang, J. Xie, M. Lin, L. Ying, et al., "Sensitivity of chest CT for COVID-19: comparison to RT-PCR," *Radiol.*, vol. 296, no. 2, pp 115-117, 2020.
- [6] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, et al., "Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases," *Radiol.*, vol. 296, no. 2, pp. 32-40, 2020.
- [7] A. Narin, C. Kaya, and Z. Pamuk, "Automatic Detection of Coronavirus Disease (COVID-19) Using X-Ray Images and Deep Convolutional Neural Networks," *arXiv*, arXiv:2003.10849, 2020.
- [8] F. Chollet, "Fundamentals of Machine Learning," in *Deep Learning with Python*, 1st ed. USA: Manning Publications Company, 2018, ch. 4, sec. 4.1, pp.93-96.
- [9] P. Patel, "Chest X-ray (Covid-19 & Pneumonia). Kaggle: Machine Learning and Data Science Community." [online]. 2020. Disponible en: <https://www.kaggle.com/prashant268/chest-xray-covid19-pneumonia?>
- [10] W. H. Khoong, "COVID-19 Xray Dataset (Train and Test Sets). Kaggle: Machine Learning and Data Science Community." [online]. 2020. Disponible en: <https://www.kaggle.com/khoongweihao/covid19-xray-dataset-train-test-sets>
- [11] T. Y. Wang, and H. M. Chiang, "Fuzzy support vector machine for multi-class text categorization," *Inf. Process. Manag.*, vol 43, no. 4, pp. 914-929, 2007.
- [12] H. Foroughi, A. Rezvanian, and A. Pazirae, "Robust Fall Detection Using Human Shape and Multi-class Support Vector Machine," in *ICVGIP'08*, Bhubaneswar, India, 2008, pp. 16-19.
- [13] M. Bouazizi, and T. Ohtsuki, "Sentiment analysis: From binary to multi-class classification: A pattern-based approach for multi-class sentiment analysis in Twitter," in *IEEE ICC 2016*, Kuala Lumpur, Malaysia, 2016, pp. 1-6.
- [14] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images," *arXiv*, arXiv:2003.09871, 2020.
- [15] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. Rajendra Acharya, (2020). "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Comput. Biol. Med.*, vol. 121, no. 103792, pp. 1-11, 2020.
- [16] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, et al., "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning." *arXiv*, arXiv:1711.05225, 2017.