

# Capítulo 4

## La inteligencia artificial y su impacto en la lucha contra la COVID-19

*Rosa María Valdovinos Rosas* <sup>1</sup>, *Angélica Guzmán Ponce* <sup>2\*</sup>, *J. Leonardo González Ruíz* <sup>1</sup>, *Javier Salas García* <sup>1</sup>

<sup>1</sup> Universidad Autónoma del Estado de México. Toluca, Estado de México, México.

[rvaldovinosr@uaemex.mx](mailto:rvaldovinosr@uaemex.mx)

[aguzman@uji.es](mailto:aguzman@uji.es)

[jlgonzalezr@uaemex.mx](mailto:jlgonzalezr@uaemex.mx)

[jsalasg@uaemex.mx](mailto:jsalasg@uaemex.mx)

<sup>2</sup> Universidad Jaime I. Castellón de la Plana, España.

[aguzman@uji.es](mailto:aguzman@uji.es)

### 4.1 Introducción

En las últimas décadas, se ha evidenciado una revolución en el campo de la medicina, con un impacto significativo de la Inteligencia Artificial

---

\*Corresponding author: [aguzman@uji.es](mailto:aguzman@uji.es)

(IA), apoyando a los profesionales médicos mediante el análisis de imágenes médicas y la asistencia en toma de decisiones clínicas [93]. Esta situación se vió más marcada con la pandemia de la COVID-19 a nivel mundial.

En diciembre de 2019, en Wuhan, provincia de Hubei, China, se registraron los primeros casos de COVID-19 causada por el virus SARS-CoV-2. Esta enfermedad se transmite principalmente a través de las gotículas que son expulsadas por una persona infectada al toser o estornudar, así como al tener contacto cercano con personas enfermas, o al tocar superficies u objetos contaminados con el virus y luego llevarse las manos a la boca, nariz u ojos.

A la fecha, se han contabilizado más de 530 millones de casos de COVID-19 y más de 6.2 millones de muertes en todo el mundo<sup>1</sup>. Como resultado del cúmulo de información generada y con el propósito de enfrentar los desafíos planteados por la pandemia, así como la explotación de los datos existentes para facilitar la comprensión del patrón de la enfermedad se han derivado una gran cantidad de estudios que utilizan algoritmos del área de IA, tales como el aprendizaje automático, reconocimiento de patrones, minería y ciencia de datos, entre otros.

La capacidad principal de la IA en el ámbito de la medicina es su habilidad para descubrir patrones en fuentes inesperadas y establecer conexiones que normalmente los humanos no podrían anticipar, así como la generación de soluciones innovadoras de una manera no convencional, lo que amplía las posibilidades de disponer de diferentes conjuntos de datos.

El presente capítulo se incluye una revisión del estado del arte en relación a la utilización de métodos de IA y su estudio con la estructura del virus SARS-CoV-2, los factores sociales asociados y el tratamiento, además del control de pacientes infectados por el virus.

---

<sup>1</sup><https://covid19.who.int/>

## 4.2 Estado del arte

Una nube de palabras ofrece una forma intuitiva de discernir la frecuencia e importancia de cada palabra, simplificando la identificación de términos clave. La nube de palabras de la Figura 4.1 se creó a partir de los resúmenes de cada artículo revisado en el estado del arte. Las palabras resaltadas son aquellas que se repiten con mayor frecuencia en comparación con las otras. Cabe mencionar que las palabras de uso frecuente que no añaden un valor significativo al texto, como los artículos determinados ('los', 'las', 'una', 'unos', etc.), así como los plurales, los tiempos verbales y los géneros, no fueron considerados.

Al analizar la nube de palabras más utilizadas en los trabajos relacionados con la IA y su impacto en la lucha contra la COVID-19, se pueden identificar predominantemente: "aprendizaje automático", "aprendizaje profundo", "redes neuronales", "análisis de datos", "modelos predictivos" y "reconocimiento de imágenes".



Figura 4.1: Nube de palabras generada de los resúmenes de trabajos relacionados.

Estas palabras reflejan la importancia de la IA en la detección temprana de casos, el monitoreo de la propagación del virus, el desarrollo de tratamientos y vacunas, así como la optimización de recursos y la

toma de decisiones informadas. La Figura 4.2 muestra los enfoques en que se puede clasificar los resultados de la revisión literaria realizada.

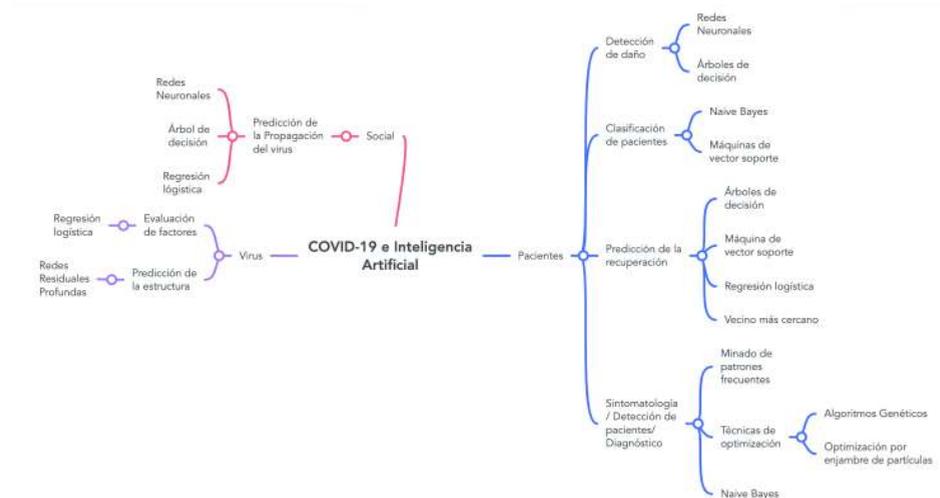


Figura 4.2: Esquema resumen de áreas de enfoque abordadas por la IA para combatir la pandemia de COVID-19.

### 4.2.1 IA y el virus SARS-CoV-2

La determinación de la estructura molecular de un virus es fundamental para comprender su funcionamiento y diseñar estrategias terapéuticas específicas. Aquí es donde la IA ha desempeñado un papel crucial al utilizar técnicas de aprendizaje automático para analizar datos estructurales en los que al utilizar información sobre virus previos, los científicos han logrado predecir con mayor precisión la estructura del nuevo virus, variantes y la creación de nuevos fármacos [45].

Senior et al. desarrollan el algoritmo AlphaFold, el cual emplea redes residuales profundas llamadas ResNets para predecir las estructuras proteicas de la proteína de membrana, la proteína 3a, nsp2, nsp4, nsp6 y el dominio C-terminal similar a la papaína del SARS-CoV-2, lo cual propicia un enorme impulso a los programas de descubrimiento de fármacos [34]. En otros trabajos como el propuesto por Pfab et

al. es propuesto un algoritmo basado en una red neuronal convolucional profunda personalizada, utilizada para derivar la estructura del complejo proteico del SARS-CoV-2 a partir de mapas de densidad de microscopía crioelectrónica de alta resolución y secuencias de aminoácidos [182].

En trabajos como el de Ong et al. [133], son aplicadas herramientas basadas en aprendizaje automático, para predecir las vacunas candidatas contra el COVID-19 basados en modelos de clasificación supervisados.

En seguimiento a los trabajos de clasificación del virus, Câmara et al [23] proponen el uso de una red neuronal convolucional (CNN) de aprendizaje profundo para clasificar y distinguir el SARS-CoV-2 de otros virus, sin limitar la longitud de la secuencia del genoma. Este enfoque logra una alta precisión al distinguir el SARS-CoV-2 de otros virus, además se sugiere que nuevos modos de representación de imagen del código genético viral, podrían mejorar la precisión de la técnica propuesta.

Raghav et al. [9] propone un método de Aprendizaje por Reforzamiento Profundo basado en el problema del bandido. El problema de bandidos se refiere a una situación donde se deben asignar secuencialmente una cantidad limitada de recursos a un conjunto fijo de acciones. Después de cada asignación, se recibe información que puede ser utilizada para tomar decisiones más informadas en rondas posteriores. De este modo, el objetivo que sigue esta investigación es optimizar la distribución de vacunas COVID-19. El enfoque está diseñado para abordar el problema de recursos limitados e inequidades de acceso. Los resultados muestran que VacSIM puede prevenir posibles infecciones y aumentar la eficacia de limitar la propagación de COVID-19 en un período de 45 días.

En seguimiento a métodos que buscan apoyar al estudio de vacunas, Muneer et al. [118] utilizaron modelos híbridos de redes neuronales profundas denominadas GCN\_GRU y GCN\_CNN, ambos surgen de la combinación de redes neuronales convolucionales de gráficos (GCN) y unidades recurrentes con compuertas (GRU) o redes neuronales convolucionales (CNN), dichos modelos pueden predecir la degradación

del Ácido ribonucleico (ARN) a partir de secuencias de ARN. El estudio demostró que modelar moléculas de ARN es crítico para comprender los mecanismos de degradación de las moléculas, lo que ayuda a minimizar problemas relacionados con la distribución de ARNm entre los pacientes. Los modelos propuestos se pueden utilizar para predecir la degradación del ARN a partir de secuencias de ARN y contribuir al desarrollo de vacunas basadas en ARNm.

Por otro lado, Elkin et al. [45] propone un enfoque de aprendizaje automático para estudiar la correlación entre las pruebas serológicas y moleculares para la infección por COVID-19, e identificar características sintomáticas significativas asociadas con la infección. El estudio recopiló datos de 2.467 donantes que fueron evaluados mediante uno o varios tipos de pruebas de COVID-19, generando características sintomáticas para representar a cada donante. Al agrupar los síntomas similares se combinaron con características demográficas, y se entrenaron cinco algoritmos de clasificación (Árboles de decisión, Regresión Logística, Máquinas de vector soporte, Redes Neuronales y refuerzo de gradientes extremo) para predecir los resultados de las pruebas de COVID-19. El estudio encontró que refuerzo de gradientes extremo logró el mejor rendimiento.

### 4.2.2 Factor social y la IA para control del virus SARS-CoV-2

En el ámbito social, la aplicación de la IA se centra en gran medida en predecir la propagación del virus SARS-CoV-2 y su impacto en la salud pública [174]. Los algoritmos de aprendizaje automático y el análisis de datos masivos han permitido a los investigadores y expertos en salud predecir cómo se propaga el virus en diferentes regiones y cómo podría evolucionar en el futuro. Estas predicciones han ayudado a los responsables de la toma de decisiones a implementar medidas preventivas y de mitigación de manera más efectiva [86].

Al respecto, Alali et al. [4] proponen un modelo basado en regresión de procesos gaussianos (GPR) y optimización bayesiana impulsada por datos para predecir con precisión la propagación de COVID-19 en

países altamente afectados como India y Brasil. El modelo tiene en cuenta la naturaleza dinámica de los datos de COVID-19 e incorpora mediciones rezagadas para mejorar su precisión. Las predicciones se encuentran dentro del intervalo de confianza del 95%.

Bhardwaj et al. [16] desarrollo de un modelo para predecir el número de casos recuperados, casos confirmados, muertes y recuento diario de casos con COVID-19. El modelo emplea Regresión Vectorial de Soporte, su principal objetivo es ayudar a los organismos gubernamentales a prepararse para otra ola y garantizar la disponibilidad de los recursos necesarios.

Por otro lado, Ayyildiz et al. [10] proponen un modelo basado en Redes Neuronales Artificiales (ANNs) para el diseño de un mecanismo efectivo de cadena de suministro de sangre para hacer frente a la pandemia de COVID-19. Primero identifican personas recuperadas para identificar posibles donantes para el tratamiento con plasma convalesciente, todo esto mediante el uso de redes neuronales, así mismo, la red permite predecir el número diario de pacientes recuperados de COVID-19. En general, el estudio concluye que el modelo propuesto es robusto y puede aplicarse ampliamente en otras partes del mundo para pronosticar el número de pacientes recuperados de COVID-19, facilitando así un mecanismo efectivo de cadena de suministro de sangre para hacer frente a la pandemia.

En línea con los avances en los métodos de predicción de la propagación del virus, Sujath et al. [176] llevaron a cabo un estudio para estimar los patrones epidemiológicos y la tasa de casos de COVID-19 en India. A través del análisis de datos recopilados de COVID-19 en Kaggle, utilizaron técnicas de regresión lineal y perceptrón multicapa para examinar los casos confirmados, las muertes y las recuperaciones en todo el país. Este enfoque permitió al modelo generar predicciones y estimaciones sobre las posibles tendencias e impactos del COVID-19 en India.

Arias y Mejía [7] utilizaron métodos de regresión lineal para analizar las infecciones históricas y pronosticadas de muertes por COVID-19. Este enfoque les permitió obtener una imagen cuantitativa de la propagación de la epidemia, incluyendo tasas de infección, letalidad y re-

cuperación en los Estados Unidos y sus ciudades y provincias desde el 21 de enero hasta el 12 de abril de 2020. Según su estudio, una estrategia de contención con cuarentenas más restrictivas podría ser más efectiva en la prevención de tasas críticas de infección y muerte.

### **4.2.3 La IA en el tratamiento y control de pacientes con COVID-19**

En la predección de la propagación del virus, el tratamiento y control de pacientes la IA también ha sido de gran importancia. Para ello, se han desarrollado métodos de diagnóstico más eficientes, que permiten una clasificación precisa de los pacientes afectados por COVID-19, así como también la predicción de su recuperación y la sintomatología y/o diagnóstico oportuno [205].

En [159] los autores buscan realizar la predicción de la propagación del virus y el desarrollo de sistemas de alerta temprana, tomado información de redes sociales, llamadas de emergencia y los sitios de noticias. El estudio permitió determinar las regiones vulnerables e identificar un grupo de casos de neumonía, prediciendo el brote de COVID-19 y la ubicación geográfica de éste con algoritmos de aprendizaje automático.

Por otro lado, Lalmauanawma et al. [88] construyeron modelos de redes neuronales artificiales y árboles de decisión a partir de imágenes de radiología o de tomografía para la detección rápida del daño en los campos pulmonares presuntivamente por el SARS-CoV-2.

Yegorov et al. [209] realizan un estudio con pacientes enfermos por COVID-19 hospitalizados en Kazajstán entre febrero y abril de 2020. Para ello se comparan los datos demográficos, clínicos, de laboratorio y radiológicos de pacientes con diferentes grados de gravedad de COVID-19 al ingreso y emplea regresión logística para evaluar los factores asociados con la gravedad de la enfermedad y la muerte hospitalaria. Los autores concluyen que la edad avanzada, las comorbilidades, el aumento del recuento de leucocitos y el sexo masculino fueron factores de riesgo para la gravedad y la mortalidad de la enfermedad COVID-19 en Kazajstán.

De igual modo, Heydari et al. [72] clasifican la información sobre los pacientes infectados por SARS-CoV-2 según las características de los pacientes en función de los síntomas clínicos, los parámetros de laboratorio y las imágenes radiológicas utilizando algoritmos de aprendizaje supervisado y no supervisado.

Los resultados obtenidos se comparan con el diagnóstico de los médicos y revelan una precisión del 92,5%, identificando como características más importantes: edad, afectación pulmonar y SPO2.

Cohen et al. presentan un modelo para predecir el grado de gravedad de la neumonía por COVID-19 utilizando imágenes de radiografía de tórax frontal para entrenar un modelo basado en redes neuronales artificiales. Los resultados obtenidos, permitieron escalar o desescalar el tratamiento otorgado, especialmente en los pacientes de cuidados intensivos [116].

El desarrollo de modelos para la predicción de la recuperación de los pacientes infectados por el COVID-19 utilizando el conjunto de datos epidemiológico de los pacientes infectados de Corea del Sur se presenta en [166]. Los autores utilizan árboles de decisión, máquina de vector soporte, regresión logística, bosque aleatorio y vecino más cercano, logrando predecir el número mínimo y máximo de días para que los pacientes con COVID-19 se recuperaran del virus, el grupo de edad de los pacientes que tienen un alto riesgo de no recuperarse y los que podrían recuperarse rápidamente. Los resultados demostraron que el modelo desarrollado es eficaz para predecir con un 99,85% de precisión la posibilidad de recuperación de los pacientes infectados.

Tandan et al. descubren patrones y reglas generales de síntomas considerando la edad, sexo, condición crónica y estado de mortalidad, entre los pacientes de COVID-19 [139]. Para ello, se analizan retrospectivamente los datos de los pacientes de COVID-19 puestos a disposición en línea por el Wolfram Data Repository. Con técnicas de aprendizaje automático y minado de patrones frecuentes, se identificaron síntomas frecuentes y se logró definir patrones en las reglas descubiertas. Al respecto, se concluyó que los síntomas más frecuente son fiebre (67%), seguido de la tos (37%), el malestar o el dolor corporal (11%), neumonía (11%) y dolor de garganta (8%). El infarto de

miocardio, la insuficiencia cardíaca y la enfermedad renal estaban presentes en menos del 1% de los pacientes. Las diez reglas de síntomas más significativos (de las 71 generadas) mostraron la tos, el shock séptico y el síndrome de dificultad respiratoria como consecuencias frecuentes.

Muhammad et al. Utilizaron regresión logística, árbol de decisión, máquina de vector soporte, Naïve Bayes y red neutral artificial utilizando un conjunto de datos epidemiológico etiquetado para casos positivos y negativos de COVID-19 en México [115]. En el estudio consideró el análisis de coeficiente de correlación entre varias características para determinar la relación existente entre ellas y de este modo desarrollar los modelos. Los resultados experimentales mostraron que el árbol de decisión brindó la mayor precisión (94,99 %), la máquina de vectores de soporte la mayor sensibilidad (93,34 %) y Naïve Bayes la mayor especificidad (94,30 %).

Shaban et al. [167] presentan una estrategia para detectar a los pacientes infectados por COVID-19, denominada Distance Biased Naïve Bayes (DBNB), de la cual se resaltan dos contribuciones: una nueva técnica de selección de características llamada Advanced Particle Swarm Optimization (APSO) que elige las características más informativas y significativas para diagnosticar a los pacientes de COVID-19 y, un nuevo modelo de clasificación, que combina pruebas de modelos de clasificación estadísticos y basados en la distancia. Los resultados experimentales demostraron que el DBNB supera a las recientes estrategias de diagnóstico de COVID-19, ya que introduce la máxima precisión con la mínima penalización de tiempo.

Mansour et al. [104] desarrollan una estrategia para apoyar el diagnóstico oportuno de COVID-19 denominada Feature Correlated Naïve Bayes (FCNB). En su funcionamiento, la propuesta consta de cuatro fases: 1. selección de características, 2. agrupación de características, 3. ponderación de características maestras y 4. la clasificación con Naïve Bayes. En la Fase 1 se seleccionan sólo las características más eficaces entre las minadas de las pruebas de laboratorio tanto para los pacientes con COVID-19 como para las personas sin el, utilizando un algoritmo genético. En la Fase 2 se construyen clusters de carac-

terísticas basados en las características seleccionadas utilizando una técnica novedosa de clustering. En la Fase 3, se asigna un peso a cada característica seleccionada utilizando un nuevo método de cálculo de pesos. En la Fase 4, se utiliza el algoritmo Naïve Bayes ponderado con modificaciones como la correlación entre características para fines de clasificación. La propuesta mostró mayor eficacia con precisiones de detección hasta del 99% respecto el estado del arte.

Con el uso de aprendizaje profundo, se han desarrollado métodos para detectar personas infectadas por COVID-19, tal es el caso del método propuesto por Hoang et al. [73], el cual detecta pacientes positivos de COVID-19 a partir de los sonidos de su tos. Dicho método consta de extraer características que representan el sonido de la tos, mediante el uso de modelos conocidos en aprendizaje profundo como (ResNet50, DenseNet121, MobileNetV1, etc.) y la caracterización a mano. Tomando estas características, evalúan y aplican algoritmos de clasificación como redes neuronales, máquinas de vector soporte y árboles de decisión. El enfoque logró el obtener resultados competitivos para el conjunto de datos Track-2 del Segundo Desafío DiCOVA 2021, obteniendo el segundo lugar.

En línea con modelos basados en redes y aprendizaje profundo, Li et al. [92] propone un modelo de diagnóstico asistido por computadora (CAD) llamado Cov-Net para el reconocimiento del COVID-19 a partir de imágenes de radiografía de tórax. El modelo utiliza una red residual modificada con convolución asimétrica y un mecanismo de atención para extraer características. Los resultados experimentales en dos bases de datos públicas de radiografía de COVID-19 demostraron la practicidad y eficacia de Cov-Net.

Con la constante creación de conjuntos de datos de rayos-x, diversos estudios han usado estos conjuntos de datos para clasificar y/o detectar el virus causado por la COVID-19, tal es el caso de Karar et al. [84] quienes proponen un marco de trabajo de clasificadores de aprendizaje profundo en cascada para mejorar el rendimiento de los sistemas de diagnóstico asistido por ordenador (CAD). Su propuesta simplifica la complicada clasificación multi-etiqueta de las imágenes de rayos X mediante una serie de clasificadores binarios para cada

caso de estado de salud examinado. La arquitectura en cascada de los clasificadores de COVID-19 y neumonía puede utilizar diferentes modelos de aprendizaje profundo afinados simultáneamente, logrando un mejor rendimiento en la confirmación de casos positivos. Los resultados sugieren un mejora en la clasificación de COVID-19, neumonía viral y neumonía bacteriana.

Siguiendo la línea de la complejidad de clasificar la COVID-19 de otras enfermedades similares, como es la neumonía, mediante el uso de conjunto de datos de rayos x, se encuentra el trabajo de Zahirul et al. [81], cuya contribución se encuentra en combinar una red neuronal convolucional (CNN) y una memoria a corto y largo plazo (LSTM, por sus siglas en inglés de long short-term memory ), este último es una red neuronal que usa unidades de memoria para almacenar información a largo plazo. Estas celdas de memoria están diseñadas para mantener y actualizar la información a través del tiempo, lo que permite a la red recordar patrones y secuencias a largo plazo. El sistema propuesto diagnositca COVID-19 con una precisión del 99.4%.

Debido a su confiabilidad y rapidez en el diagnóstico, diversos estudios hacen uso de la tomografía computarizada (TC) de tórax, tal es el caso de Shah et al. [168], los cuales desarrollaron un método basado en aprendizaje profundo con redes como CTnet-10, DenseNet-169, VGG-16, ResNet-50, InceptionV3 y VGG-19 para el diagnóstico temprano y preciso de COVID-19. Los resultados mostraron que el modelo VGG-19 logró la mayor precisión.

Por otro lado, Wang et al. [200] propone el uso de la red UNet para el diagnóstico de COVID-19. La UNet es una red profunda en forma de U, que consiste en una capa de codificación y una capa de decodificación simétricas conectadas por un puente o paso de información de menor resolución, que ayuda a preservar la información espacial en la segmentación de imágenes. La capa de codificación reduce la resolución espacial de la imagen y extrae características, mientras que la capa de decodificación aumenta la resolución espacial y utiliza información de características para generar una máscara de segmentación. Por lo que, el modelo usa la UNet para segmentar la región del pulmón, posteriormente usa una red neuronal profunda 3D para prede-

cir la probabilidad de infección por COVID-19 y localizar lesiones de COVID-19.

## 4.3 Conclusiones

En este capítulo se realizó la revisión bibliográfica de algunos de los estudios que contribuyen alternativas de solución para el diagnóstico, análisis y tratamiento del COVID-19.

De los trabajos revisados, la mayoría de ellos se centran en identificar y estudiar los factores de riesgo que influyen en la aparición de COVID-19 en pacientes, en tanto que poco menos se centran en el análisis y propuesta de tratamientos certeros una vez adquirida la enfermedad.

Respecto a los algoritmos más utilizados, se pudo observar que la mayoría de los trabajos hacen uso de algoritmos de clasificación como árboles de decisión k-NN y Naïve Bayes. En relación a las técnicas descriptivas para realizar el minado de datos, la mayoría hace uso de algoritmos de agrupamiento y otras técnicas de selección y extracción de características para la detección y el diagnóstico. Algo interesante de observar es la poca utilización de los métodos de aprendizaje por refuerzo, clasificación híbrida o técnicas de optimización, como los algoritmos genéticos y la optimización de enjambre de partículas.

Producto de los resultados experimentales de las investigaciones anteriormente expuestas indicaron que la edad y los síntomas son los dos factores dominantes en el desenlace de los pacientes infectados. De igual modo, los factores dominantes coincidentes en los trabajos analizados, que influyen en posibles complicaciones de la enfermedad, indican tres características principales: la gravedad de la enfermedad, la edad del paciente y el historial de enfermedades crónicas.

La mayoría de los trabajos existentes en la literatura están más centrados en identificar la COVID-19 en pacientes. Sin embargo, la aplicación de técnicas de IA en el análisis de tratamientos certeros una vez adquirida la enfermedad también es esencial.

Independientemente del creciente número de estudios que se desarrollan alrededor del mundo, aún existen áreas de estudios por explorar, tales como la predicción de rebrotes de la enfermedad, la identificación y diferenciación entre las variantes de cepas del SARS-CoV-2 y otras formas de neumonía, efectos secundarios de las vacunas, estimación de la inmunidad adquirida por vacunas o de forma natural, entre otros. Estos estudios ayudarán a especialistas del área médica al rededor del mundo a tratar de forma más certera la enfermedad, disminuyendo la aparición de secuelas y sobretodo disminuir las muertes a causa de la enfermedad o a consecuencia de ella.

Por último, los posibles escenarios a corto y mediano plazo, se orientan a vislumbrar cómo la IA puede influir en las siguientes pandemias, de tal modo que los estudios existentes hasta el día de hoy podrían permitir su utilización para prepararnos mejor, transformándose de este modo en una herramienta de prevención y no solamente de ayuda a la solución del problema.